Firm Characteristics and the Cross-Section of Covariance Risk☆

Chris Kirby[a]

[a] *UNC Charlotte, 9201 University City Blvd., Charlotte, NC 28223; phone: (704) 687-0845; email: ckirby10@uncc.edu*

Abstract

I analyze the cross-section of covariance risk for individual stocks using a new type of multivariate volatility model in which firm characteristics serve as time-varying loadings on fundamental factors. The evidence points to strong linkages between firm characteristics and covariance risk, and also reveals that cross-sectional differences in covariance risk explain much of the cross-sectional variation in expected excess stock returns. I find, for example, that the fundamental factors perform at least as well as the Fama-French factors in regression-based pricing tests. In view of its tractability and performance, the proposed model should find use in a variety of applications.

Keywords: fundamental factor, multivariate GARCH, cross-section of expected returns, price of covariance risk, Fama-MacBeth regressions

---

☆Address correspondence to: Chris Kirby, Belk College of Business, University of North Carolina at Charlotte, 9201 University City Blvd., Charlotte, NC 28223, phone: (704) 687-0845, email: ckirby10@uncc.edu.

1. Introduction

Although researchers have known for decades that firm characteristics help to explain the cross-section of average stock returns, many still question whether this finding can be reconciled with the predictions of asset pricing theory. Among those who advance rational pricing stories, it is commonly argued that characteristics proxy for exposures to systematic risk. I assess the ability of firm characteristics to capture systematic risk using a new type of multivariate generalized autoregressive conditional heterskedasticity (MGARCH) model that assumes a fundamental factor structure for individual stock returns. Because the features of the model make it feasible to estimate the conditional covariance matrix of returns for systems containing thousands of individual stocks, the analysis delivers a comprehensive picture of the relation between firm characteristics, the cross-section of covariance risk, and the cross-section of average stock returns.

My approach to model development builds on ideas pioneered by Rosenberg (1974) and Fama and French (1993). Following Rosenberg (1974), I assume that the innovations to individual stock returns have a common factor structure in which firm characteristics function as observable factor loadings. To eliminate any issues with stationarity and aid in interpretation, I assume that each of the time-varying characteristics is standardized to have a cross-sectional mean of zero and a cross-sectional variance of one in every time period. A major departure from typical conditional factor specifications is that the innovations to individual stock returns are linked to the factor innovations and idiosyncratic errors via a full-rank matrix transformation that entails no loss of information. As a result, the factor innovations are indentified with characteristic-based hedge portfolios that have known time-varying weights. Fama and French (1993) introduced the use characteristic-based portfolio returns as priced factors. In their case, however, the portfolio weights are prespecified rather than emerging naturally from the underlying assumptions of the model.

Under the proposed factor structure, the conditional covariances between individual stock returns can be expressed in terms of the conditional covariances between the factors. I assume that the conditional factor covariance matrix displays MGARCH dynamics, and refer to the resulting specification for individual stock returns as a fundamental-factor MGARCH (FF-MGARCH) model. The FF-MGARCH model has some structural elements in common with the generalized orthogonal GARCH model of Lanne and Saikkonen (2007). However, it is far more computationally tractable because the factor loadings are observable. This makes likelihood-based inference feasible for systems that contain thousands of individual stocks. Importantly, missing values and time-series changes in the number of traded stocks pose no difficulties whatsoever. To incorporate and test the predictions of asset pricing theory, I specify the vector of conditional means such that the model implies exact factor pricing. This is accomplished by assuming that conditional expected excess returns are linearly related to the conditional covariances between the excess returns and fundamental factors. The price of covariance risk for each factor is assumed to be constant.

The choice of characteristics determines the nature of the fundamental factors. I assume that the first characteristic has a value of one for every firm, and show that the associated fundamental factor is simply the excess return on the equally-weighted market portfolio. The next four characteristics might be described as the "usual suspects" from the empirical asset pricing literature: the logarithm of market equity (log ME), the logarithm of the book-to-market equity ratio (log BE/ME), gross profitability scaled by book assets (GP/BA), and the logarithm of the gross growth rate of book assets (log AG). I also include two variables that have been widely studied in the anomalies literature: current accruals scaled by book assets (CA/BA), and the logarithm of the gross growth rate of split-adjusted shares outstanding (log SG). Finally, I consider net sales scaled by market equity (NS/ME). This characteristic has received little attention outside of a recent study by Lewellen (2015), but it appears to have incremental explanatory power for cross-section of average returns, particularly for small-cap stocks. I refer to the returns on the associated hedge portfolios as the size, value, profitability, investment, accruals, offerings, and sales factors, respectively.

I fit the FF-MGARCH model to monthly individual excess stock returns for NYSE, NASDAQ, and AMEX firms. Because I include every stock that has the required data items each month, the cross-sectional dimension of the dataset changes through time. It contains around 3500 stocks on average for my sample period (July 1962 to December 2016). I also investigate the impact of excluding microcap stocks on my findings. This is a common robustness check in the empirical literature on characteristic-based asset pricing models. The parameter estimates are obtained by maximizing the log likelihood function for conditionally Gaussian observations. However, I employ a multi-step estimation procedure of the type often used in MGARCH studies. This significantly reduces the computational demands of fitting the model.

As anticipated, the evidence with respect to second moment dynamics is consistent with that from the volatility modeling literature. The parameter estimates point to strong persistence in the conditional factor variances, conditional factor correlations, and conditional error variances. This finding, in conjunction with the strict factor structure that underpins the model, suggests that the conditional second moments of individual excess stock returns are subject to long swings away from their long-run average values. The parameter estimates for the no-microcaps sample display the same general patterns. Although the estimated persistence measures are somewhat higher after dropping microcap stocks, it is apparent that the second moment dynamics of the factors extracted from the no-microcaps sample are broadly similar to those extracted from the full sample.

Examining the distributional properties of the estimated conditional covariances of excess stock returns with the fundamental factors yields some initial insights on the cross-sectional implications of the model. Notably, the estimated conditional covariances of excess stock returns with the market factor vary widely across firms, suggesting that the cross-sectional variation in conditional market betas is substantial. This is an interesting finding because the loading on the market factor is the same for every firm. Thus it follows that the cross-sectional variation in estimated market betas arises from the cross-sectional variation in the characteristics in conjunction with the estimated conditional correlations between the market factor and the other factors in the model. In other words, the cross-sectional variation in the estimated market betas is driven entirely by characteristic-based patterns in the data. Stocks with high estimated betas have excess returns that display low (often negative) conditional covariances with the size, value, and profitability factors, and high conditional covariances with the investment, accruals, issuance, and sales factors.

Under the model, every factor except the market contributes to differences in systematic risk across firms. As might be inferred from the properties of the estimated market betas, the evidence suggests that firms display substantial variation in their risk exposures. The size and value factors stand out in this regard, with each exhibiting a wide range of estimated conditional covariances with excess stock returns. This finding is consistent with the strong incremental explanatory power of the SMB and HML returns in the Fama and French (1993) generalization of the Sharpe (1964)–Lintner (1965) capital asset pricing model (CAPM). The range of estimated conditional covariances for the remaining five factors is narrower, which is suggestive of a smaller role in explaining cross-sectional variation in systematic risk, but the results suggest that all seven factors make non-negligible contributions in this regard. In short, the findings are consistent with the presence of strong linkages between the firm characteristics and covariance risk, which supports the use of the FF-MGARCH specification as a risk model in high-dimensional applications.

The performance of the FF-MGARCH specification as an asset pricing model is a separate issue. Broadly speaking, the evidence on the pricing ability of the fundamental factors is mixed. The estimated price of covariance risk is statistically significant for every factor, which indicates that the conditional covariances of excess returns with the factors help to explain the differences in conditional expected excess stock returns across firms. Hence, I find support for the hypothesis that the fundamental factors are "priced" in the parlance of the traditional two-pass regression-based approach to testing asset pricing models. This finding is in line with the predictions of exact factor pricing. However, an analysis of the pricing errors makes it clear that the model falls short of fully capturing the cross-section of conditional expected excess stock returns.

To keep the analysis manageable and facilitate comparisons with prior research, I use well-diversified portfolios to evaluate the pricing performance of the model. Specifically, I employ portfolios that are formed by using the fitted conditional expected excess stock returns to sort stocks into 25 groups for each month in the sample period. The idea is similar to the well-established practice of using portfolio sorts to illustrate the extent to which conditioning on firm characteristics spreads average stock returns. To see how well the model captures differences in unconditional expected stock returns across firms, I compare the average excess portfolio returns to the average values of the fitted conditional expected excess portfolio returns.

The 25 portfolios display a wide spread of average annualized excess returns: -4.2% to 35.9%. This is indicative of the extent to which the covariances of excess returns with the fundamental factors capture the characteristic-based patterns in the data. In comparison, the average values of the fitted conditional expected excess portfolio returns range from -10.9% to 31.7%. As these figures suggest, the average estimated pricing error is slightly negative. Portfolios that fall near the ends of the range of estimated expected excess returns have positive estimated pricing errors, while those that fall in the center of this range have negative estimated pricing errors. The median and average of the absolute estimated pricing errors are 2.0% and 2.6% per annum, respectively. Estimated pricing errors of this magnitude seem likely to be economically significant.

To investigate further, I examine the pricing performance of the fundamental factors in the arbitrage pricing theory (APT) framework of Ross (1976). This facilitates head-to-head pricing comparisons with competing models from the literature. I use the five-factor model of Fama and French (2015) as a performance benchmark. In the APT framework, the estimated pricing errors are simply the estimated intercepts in regressions of excess portfolio returns on a constant and the factors. I find that most of the estimated intercepts for fundamental factor regressions are statistically significant, indicating that the hypothesis of exact factor pricing is rejected. This bolsters the conclusion that conditional covariances of excess stock returns with fundamental factors fail to fully explain the cross-section of conditional expected excess stock returns.

In comparison, the five-factor model produces a smaller number of statistically-significant intercepts. This seems unfavorable to the FF-MGARCH model at first glance. However, the reduction in statistical significance relative to the fundamental factor regressions is driven by increases in the standard errors. The average magnitude of the estimated intercepts is 2.6% per annum for the fundamental factors, versus 3.5% per annum for the Fama and French (2015) factors. The fundamental factors produce lower standard errors because they explain more of the time series variation in the excess portfolio returns than the Fama and French (2015) factors. On average, the regression R^2 is 94.0% for the former versus 85.4% for the latter. Hence, the general picture that emerges from the comparisons casts the FF-MGARCH model in a relatively favorable light.

Overall the analysis suggests that the FF-MGARCH model is a very promising addition to the small set of existing MGARCH models that are designed to capture time-varying covariances in high-dimensional settings. Although there are certainly indications of specification error, this is hardly surprising. One would be hard pressed to argue that a parsimoniously-parameterized MGARCH model can be expected to unerringly capture the dynamics of excess returns for thousands of individual stocks. The real question is not whether the model is misspecified, but whether it provides a reasonably accurate description of the process that generates excess stock returns. Because all indications are that the FF-MGARCH model enjoys considerable success in this regard, I anticipate that it will prove useful in a wide variety of applications.

2. Fundamental-Factor MGARCH Model for Individual Stock Returns

For many years, the use of MGARCH models was almost entirely confined to low-dimensional settings because the parameter space of most models grows very quickly with the number of variables in the system. Researchers have recently developed a few tightly-parameterized MGARCH models, such as the dynamic equicorrelation (DECO) specification of Engle and Kelly (2012), that can be applied to high-dimensional systems, and MGARCH models based on factor structures, such as those proposed by Alexander (2001), van der Weide (2002), Lanne and Saikkonen (2007), and Fan et al. (2008), show promise in balancing the competing demands of generality and computational tractability. However, none of these models is suited to the task of analyzing highly-unbalanced panels of stock returns for hundreds or thousands of individual firms, many of which appear in the dataset for a relatively short window of time.

To develop an MGARCH model that is specifically designed for investigating the relation between characteristics and covariance risk, I build on ideas pioneered by Rosenberg (1974) and Fama and French (1993). Rosenberg (1974) extends the familiar market model by allowing the beta of each stock to be linearly related to an observable set of firm characteristics. This extension yields what is commonly called a fundamental factor model: a factor model in which characteristics function as observable factor loadings. Fama and French (1993), on the other hand, investigate the asset pricing performance of a model that has two characteristic-based factors. Specifically, they use the returns on two characteristic-based hedge portfolios to formulate a three-factor generalization of the Sharpe (1964)–Lintner (1965) CAPM. In both cases, the modeling strategy is developed with an eye towards capturing the observed characteristic-based patterns in average stock returns.

I follow Rosenberg (1974) by assuming that firm characteristics can be treated as observable factor loadings. This assumption lays the foundation for developing a factor-based MGARCH model whose computational demands rise very slowly with the dimensionality of the system. Notably, missing values for individual stock returns and time-series changes in the number of traded stocks do not pose any computational difficulties in the proposed framework. This is because the model implies that the factors are returns on characteristic-based hedge portfolios that are rebalanced in every time period. Although period-by-period rebalancing of the hedge portfolios is in the spirit of Fama and French (1993), the approach used to construct the hedge portfolios relies on cross-sectional regression methods instead of an *ad hoc* characteristic-based sorting scheme.

2.1. Common factor structure for returns

I begin by describing the factor structure that underpins the model. Let $\boldsymbol{r}_{t+1}$ denote an $N \times 1$ vector of excess stock returns for period $t+1$. For example, it might represent the vector of excess returns for all NYSE, NASDAQ, and AMEX stocks for a particular month. Let $\mathcal{I}_t$ denote the information set of market participants for period t. I take as given that $\boldsymbol{r}_{t+1}$ can be expressed as

$$\boldsymbol{r}_{t+1} = \boldsymbol{m}_t + \boldsymbol{u}_{t+1}, \tag{1}$$

where $\boldsymbol{m}_t = \mathrm{E}(\boldsymbol{r}_{t+1}|\mathcal{I}_t)$ denotes the $N \times 1$ vector of conditional expected excess returns and $\boldsymbol{u}_{t+1}$ is an $N \times 1$ vector of serially-independent innovations. The proposed dynamic specification for $\boldsymbol{S}_t = \mathrm{E}(\boldsymbol{u}_{t+1}\boldsymbol{u}'_{t+1}|\mathcal{I}_t)$, the conditional covariance matrix of $\boldsymbol{r}_{t+1}$, builds on two key assumptions.

First, I assume that the vector of serially-independent innovations in equation (1) has a decomposition of the form

$$\boldsymbol{u}_{t+1} = \boldsymbol{B}_t(\boldsymbol{f}_{t+1} - \boldsymbol{m}_{f,t}) + \boldsymbol{G}_t\boldsymbol{e}_{t+1}, \tag{2}$$

where $\boldsymbol{B}_t \in \mathcal{I}_t$ is an $N \times K$ matrix, $\boldsymbol{G}_t \in \mathcal{I}_t$ is an $N \times (N-K)$ matrix, $\boldsymbol{f}_{t+1}$ is a $K \times 1$ vector of common factors whose conditional mean is $\boldsymbol{m}_{f,t} = \mathrm{E}(\boldsymbol{f}_{t+1}|\mathcal{I}_t)$, and $\boldsymbol{e}_{t+1}$ is an $(N-K) \times 1$ vector of errors that satisfies $\mathrm{E}(\boldsymbol{e}_{t+1}|\mathcal{I}_t) = \boldsymbol{0}$, $\mathrm{E}(\boldsymbol{e}_{t+1}\boldsymbol{f}'_{t+1}|\mathcal{I}_t) = \boldsymbol{0}$, and $\mathrm{E}(\boldsymbol{e}_{t+1}\boldsymbol{e}'_{t+1}|\mathcal{I}_t) = d_t\boldsymbol{I}$. Equation (2) implies that $\boldsymbol{u}_{t+1}$ is described by a type of linear factor model that permits the matrix of factor loadings and the conditional covariance matrix of $\boldsymbol{G}_t\boldsymbol{e}_{t+1}$, the $N \times 1$ vector of idiosyncratic return shocks, to change through time. The main departure from typical specifications of conditional factor models is that the leading dimension of $\boldsymbol{e}_{t+1}$ is smaller than that of $\boldsymbol{u}_{t+1}$, so the conditional covariance matrix of the vector of idiosyncratic shocks is singular. This feature reflects the nature of the factor innovations in the model. If the matrix $\boldsymbol{A}_t = (\boldsymbol{B}_t, \boldsymbol{G}_t)$ is nonsingular, then the factor innovations are simply linear combinations of the demeaned excess stock returns. Note in particular that $\boldsymbol{f}_{t+1} - \boldsymbol{m}_{f,t}$ is given by the first K elements of the $N \times 1$ vector $\boldsymbol{A}_t^{-1}\boldsymbol{u}_{t+1}$.

Second, I assume that conditional factor loadings are observable firm characteristics, and hence the nth row of $\boldsymbol{B}_t$ contains the observed characteristic values of the nth firm for period t. Under this assumption, the columns of $\boldsymbol{G}_t$ form a basis for the null space of an observable projection matrix. Specifically, $\boldsymbol{G}_t$ contains the first $N-K$ eigenvectors of $\boldsymbol{M}_t = \boldsymbol{I} - \boldsymbol{B}_t(\boldsymbol{B}'_t\boldsymbol{B}_t)^{-1}\boldsymbol{B}'_t$ in its columns. This follows because the inverse of $\boldsymbol{A}_t$ is given by

$$\boldsymbol{A}_t^{-1} = \begin{pmatrix} (\boldsymbol{B}'_t\boldsymbol{B}_t)^{-1}\boldsymbol{B}'_t \\ \boldsymbol{G}'_t \end{pmatrix} \tag{3}$$

for any choice of $\boldsymbol{B}_t$ that has full column rank.[1]

In view of equation (3), it is easy to see that the vector of fundamental factors for period $t+1$ can be expressed as

$$\boldsymbol{f}_{t+1} = \boldsymbol{m}_{f,t} + \boldsymbol{v}_{t+1}, \tag{4}$$

where $\boldsymbol{v}_{t+1} = (\boldsymbol{B}_t'\boldsymbol{B}_t)^{-1}\boldsymbol{B}_t'\boldsymbol{u}_{t+1}$ is the vector of fundamental factor innovations. Notice that the $K \times N$ matrix $\boldsymbol{B}_t^+ = (\boldsymbol{B}_t'\boldsymbol{B}_t)^{-1}\boldsymbol{B}_t'$ takes a familiar form: the left pseudoinverse of $\boldsymbol{B}_t$ that delivers the ordinary least squares (OLS) estimator of the coefficients for a cross-sectional regression of $\boldsymbol{u}_{t+1}$ on $\boldsymbol{B}_t$. Hence, we can view equation (2) as a regression-based decomposition of $\boldsymbol{u}_{t+1}$ in which the regression residuals $\boldsymbol{e}_{t+1}$ are assumed to satisfy a set of conditional orthogonality restrictions. Because this decomposition lies at the core of my modeling strategy, the proposed specification for $\boldsymbol{S}_t$ inherits many of the features that have made cross-sectional regressions a workhorse of the empirical asset pricing literature. For example, allowing N to change from one period to the next poses no difficulties whatsoever. This is essential given the intended application of the model.

Lanne and Saikkonen (2007) develop a generalized orthogonal GARCH model that is based on a linear decomposition similar to that in equation (2). However, their model assumes that the matrix $\boldsymbol{A}_t$ is both time invariant and unobservable. The potential gains from taking this matrix to be a known function of observable time-varying characteristics are clear. First, we allow the conditional factor loadings to change through time. Second, we drastically reduce parameter proliferation as the value of N increases. Third, we make likelihood-based inference feasible for high-dimensional problems. This is because the computational demands of fitting a specification for $\boldsymbol{S}_t$ are not very sensitive to the value of N for the typical case in which the joint conditional distribution of excess returns is assumed to be multivariate normal.

The assumption that $\mathrm{E}(\boldsymbol{e}_{t+1}\boldsymbol{e}_{t+1}'|\mathcal{I}_t) = d_t\boldsymbol{I}$ may appear to be unduly restrictive. But similar assumptions appear in the literature on static factor analysis. For instance, the probabilistic version of principal components analysis is based on a static factor model in which the covariance matrix of the errors is assumed to be a scalar multiple of the identity matrix (Tipping and Bishop, 1999).

[1] To see how the inverse is derived, note that $\boldsymbol{M}_t$ is symmetric, idempotent, and has rank $N-K$. Because these properties imply that its eigenvalues consist of $N-K$ ones and K zeros, it can be decomposed as $\boldsymbol{M}_t = \boldsymbol{G}_t\boldsymbol{G}_t'$, where $\boldsymbol{G}_t$ is an $N \times (N-K)$ matrix with orthonormal columns, i.e., $\boldsymbol{G}_t'\boldsymbol{G}_t = \boldsymbol{I}$. Using these results along with $\boldsymbol{M}_t\boldsymbol{B}_t = \boldsymbol{0}$, it follows that $\boldsymbol{G}_t'\boldsymbol{B}_t = \boldsymbol{0}$. Thus it is easy to verify that the matrix in equation (3) satisfies $\boldsymbol{A}_t^{-1}\boldsymbol{A}_t = \boldsymbol{A}_t\boldsymbol{A}_t^{-1} = \boldsymbol{I}$.

The success of such methods in a range of settings suggests that the assumption of spherical errors should be a reasonable starting point for model development. Of course additional flexibility can be introduced at the cost of increased computational demands. One could, for example, replace the assumption of spherical errors with $\mathrm{E}(\boldsymbol{e}_{t+1}\boldsymbol{e}'_{t+1}|\mathcal{I}_t) = d_t\boldsymbol{I} + c_t(\mathbf{1}\mathbf{1}' - \boldsymbol{I})$, where $d_t > c_t > 0$ and $\mathbf{1}$ denotes an $(N-K)\times 1$ vector of ones. This would yield a conditional factor model for $\boldsymbol{u}_{t+1}$ in which the errors display time-varying equicorrelation.

2.2. A brief digression on the characteristics-versus-covariances debate

Before filling in the details of the model, it is useful to discuss how specifying characteristics as factor loadings fits into the characteristics-versus-covariances debate. Numerous studies show that firm characteristics help to explain the cross-section of average stock returns. However, there is no consensus view on the explanation for this finding. Some researchers, such as Fama and French (1993, 1996, 2000), favor rational pricing stories. They argue that firm characteristics, such as market capitalization and the book-to-market ratio, capture cross-sectional differences in expected stock returns by serving as proxies for the covariances between returns and common risk factors. Others, such as Daniel and Titman (1997, 1998), favor behavioral stories. They argue that the covariances between returns and common risk factors provide little information about expected stock returns after controlling for firm characteristics. Interestingly, fundamental factor structures have the potential to explain the empirical findings of researchers on both sides of the debate.

For instance, many of the findings deal with the pricing performance of the Fama and French (1993) three-factor model. Under this model, the systematic risk of a stock depends on how its excess return covaries with three risk factors: the excess return on the value-weighted market portfolio (the VWM portfolio), the return on a hedge portfolio that short sells large-cap stocks to purchase small-cap stocks (the SMB portfolio), and the return on a hedge portfolio that short sells stocks with low B/M ratios to purchase stocks with high B/M ratios (the HML portfolio). Fama and French (1993) show that the estimated loadings on these factors successfully capture the patterns in average excess returns for portfolios that are formed by sorting stocks on market capitalization and the book-to-market ratio. This lines up with what we would expect to find under a fundamental factor model in which market capitalization and the book-to-market ratio are factor loadings. That is, we would expect the estimated slope coefficients obtained by regressing excess stock returns on the fundamental factors, which are returns on characteristic-based hedge portfolios, to do a reasonable job of capturing the characteristic-based patterns in average excess stock returns.

Daniel and Titman (1997) also investigate the explanatory power of the estimated loadings on the SMB and HML factors. However, their tests center on sets of portfolios whose constituent stocks have roughly the same market capitalization and same book-to-market ratio. They report that there is no apparent relation between the average excess portfolio returns and the estimated slopes on the SMB and HML factors, and conclude that "it is characteristics rather than factor loadings that determine expected returns." Once again, this what we would expect to find under a fundamental factor model in which market capitalization and the book-to-market ratio are factor loadings. The variation in the estimated loadings for stocks or portfolios that have the same values of the characteristics must be due to estimation error. Thus differences in the estimated loadings should not have any ability to explain differences in the average excess stock or portfolio returns.

More broadly, it is easy to envision other scenarios in which such tests are unlikely to be fruitful. Suppose, for example, that excess stock returns are described by a conditional linear factor model, the unobserved time-varying loadings are cross-sectionally correlated with firm characteristics, and exact factor pricing holds. If we use the characteristics as loadings and extract the associated factors, then the estimated factor loadings (estimated regression slopes) will have some ability to explain the cross-section of average excess stock returns. However, the explanatory power of the characteristics will likely dominate that of the estimated loadings, because the fundamental factors maximize the cross-sectional explanatory power of the characteristics for individual excess stock returns.

Instead of trying to determine whether covariances proxy for characteristics or vice versa, I take an in-depth look at the pricing implications of fundamental factor models. The basic idea is to specify $\boldsymbol{m}_t$ such that exact factor pricing holds, use a flexible multivariate GARCH process to capture the dynamics of the conditional factor covariance matrix, and estimate the price of covariance risk for each factor by fitting the model to excess returns for large universe of individual stocks. My objectives are to assess the extent to which firm characteristics explain the cross-section of covariance risk, assess the extent to which the covariances with the fundamental factors explain the cross-section of expected excess returns, and provide insights on the relative importance of the different fundamental factors in determining the overall pricing performance of the model.

2.3. Modeling the cross-section of conditional expected excess stock returns

Equations (1) and (2) deliver exact factor pricing if the cross-sectional variation in the values of the characteristics explains all of the cross-sectional variation in conditional expected excess stock returns. By specifying $\boldsymbol{m}_t$ in accordance with this restriction,

$$\boldsymbol{m}_t = \boldsymbol{B}_t \boldsymbol{m}_{f,t}, \tag{5}$$

I can obtain evidence on the pricing performance of the fundamental factors. To lay the groundwork for the proposed specification of $\boldsymbol{m}_{f,t}$, I invoke two additional assumptions about the nature of the characteristics that serve as conditional factor loadings in the model.

First, I assume that the leading column of $\boldsymbol{B}_t$ is an $N \times 1$ vector of ones. This is analogous to including an intercept in a cross-sectional regression model. It allows for the presence of a cross-sectionally invariant component in conditional expected excess stock returns. Second, I assume that the remaining $K-1$ characteristics have been standardized such that $(1/N)\sum_{n=1}^{N} b_{nk,t} = 0$ and $(1/N)\sum_{n=1}^{N} b_{nk,t}^2 = 1$, where $b_{nk,t}$ denotes the kth element of the nth row of $\boldsymbol{B}_t$. Using standardized characteristics makes it reasonable to treat the matrix of conditional factor loadings as stationary, and also makes it easier to compare the estimates of the parameters associated with different characteristics. Both assumptions are easily satisfied in empirical work. Furthermore, they identify the first element of $\boldsymbol{f}_{t+1}$ as the excess return on the equally-weighted market portfolio for period $t+1$.[2] This follows by noting that $\boldsymbol{B}_t'\boldsymbol{B}_t$ is a block diagonal matrix whose first row is $(N, 0, \ldots, 0)$, and hence the first row of $\boldsymbol{B}_t^{+}$ is $(1/N, \ldots, 1/N)$. Isolating the market factor as a distinct element of $\boldsymbol{f}_{t+1}$ makes the pricing implications of the model more transparent.

Consider, for example, the specification $\boldsymbol{m}_{f,t} = \boldsymbol{H}_t\boldsymbol{\lambda}$, where $\boldsymbol{H}_t = \mathrm{E}(\boldsymbol{v}_{t+1}\boldsymbol{v}_{t+1}'|\mathcal{I}_t)$ denotes the conditional covariance matrix of $\boldsymbol{f}_{t+1}$. Substituting for $\boldsymbol{m}_{f,t}$ in equation (5) yields

$$\boldsymbol{m}_t = \boldsymbol{C}_t\boldsymbol{\lambda}, \tag{6}$$

where $\boldsymbol{C}_t = \boldsymbol{B}_t\boldsymbol{H}_t$ denotes the $N \times K$ matrix of conditional covariances between $\boldsymbol{r}_{t+1}$ and $\boldsymbol{f}_{t+1}$.

[2]In other words, the first element of $\boldsymbol{m}_{f,t} = \boldsymbol{B}_t^{+}\boldsymbol{m}_t$ is the conditional expected excess return on an equally-weighted portfolio of the N stocks, and the first element of $\boldsymbol{v}_{t+1} = \boldsymbol{B}_t^{+}\boldsymbol{u}_{t+1}$ is the corresponding return innovation.

Hence, specifying $\boldsymbol{m}_{f,t} = \boldsymbol{H}_t\boldsymbol{\lambda}$ implies that the conditional price of covariance risk is constant for each of the K factors. Adopting a constant-price-of-risk specification seems like a straightforward approach for developing insights about the cross-section of covariance risk captured by the fundamental factors. However, one of its pricing implications is not very palatable.

Note that the first column of $\boldsymbol{C}_t$ is the vector of conditional covariances of excess stock returns with the market factor. Each element of this vector is a characteristic-weighted sum of the elements in the first column of $\boldsymbol{H}_t$. Thus the covariances will differ across firms unless the excess return on the market portfolio is conditionally uncorrelated with all of the remaining factors. Because equation (6) treats the conditional covariance with the market as a separate source of covariance risk, it implies that cross-sectional differences in exposure to this risk make a distinct marginal contribution to cross-sectional differences in conditional expected excess stock returns. This seems untenable given that the cross-sectional differences in the conditional covariance with the market arise solely from the conditional correlations of the market with the other factors.

To eliminate this concern, I employ a modified version of the constant-price-of-risk specification for the empirical analysis. Let $\boldsymbol{H}_t$ be partitioned as

$$\boldsymbol{H}_t = \begin{pmatrix} h_{11,t} & \boldsymbol{h}_{12,t} \\ \boldsymbol{h}_{21,t} & \boldsymbol{H}_{22,t} \end{pmatrix} \tag{7}$$

where $h_{11,t}$ is a scalar, $\boldsymbol{h}_{12,t}$ and $\boldsymbol{h}_{21,t}$ are row and column vectors with $K-1$ elements, and $\boldsymbol{H}_{22,t}$ is a $(K-1)\times(K-1)$ matrix. Similarly, partition $\boldsymbol{f}_{t+1} = (f_{1,t+1}, \boldsymbol{f}_{2,t+1})'$ and $\boldsymbol{\lambda} = (\lambda_1, \boldsymbol{\lambda}_2)'$ in a conformable fashion. I adopt the specification $\boldsymbol{m}_{f,t} = \boldsymbol{H}_t\boldsymbol{L}_t\boldsymbol{\lambda}$, where $\boldsymbol{L}_t$ is given by

$$\boldsymbol{L}_t = \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{H}_{22,t}^{-1}\boldsymbol{h}_{21,t} & \boldsymbol{I} \end{pmatrix}. \tag{8}$$

The motivation for this approach, which is equivalent to specifying $\boldsymbol{\lambda}_t = \boldsymbol{L}_t\boldsymbol{\lambda}$ as the price of risk vector for period $t+1$, may not be obvious at first glance. With a little algebra, however, the resulting specification for $\boldsymbol{m}_t$ can be expressed as

$$\boldsymbol{m}_t = \boldsymbol{C}_t^*\boldsymbol{\lambda}, \tag{9}$$

where $\boldsymbol{C}_t^* = \boldsymbol{B}_t \boldsymbol{H}_t \boldsymbol{L}_t$ is obtained by replacing the first column of $\boldsymbol{C}_t$ with $(h_{11,t} - \boldsymbol{h}_{12,t}\boldsymbol{H}_{22,t}^{-1}\boldsymbol{h}_{21,t})\mathbf{1}$.

Equation (9) has a simple interpretation. Let $f_{1,t+1}^*$ denote the component of the market factor that is conditionally uncorrelated with the other factors, i.e.,

$$f_{1,t+1}^* = f_{1,t+1} - \boldsymbol{f}_{2,t+1}' \boldsymbol{H}_{22,t}^{-1} \boldsymbol{h}_{21,t}. \tag{10}$$

It is easy to verify that

$$\mathrm{var}(f_{1,t+1}^* | \mathcal{I}_t) = h_{11,t} - \boldsymbol{h}_{12,t}\boldsymbol{H}_{22,t}^{-1}\boldsymbol{h}_{21,t}, \tag{11}$$

which shows that λ_1 represents the price of *variance risk* for an orthogonalized version of the market factor. The conditional variance of $f_{1,t+1}^*$ is priced because it represents the unavoidable component of market risk. To see why it is unavoidable, note that the model implies that $\mathrm{cov}(r_{n,t+1}, f_{1,t+1}^* | \mathcal{I}_t) = \mathrm{var}(f_{1,t+1}^* | \mathcal{I}_t)$ for all n. So every stock is exposed to the same baseline level of variance risk, and the compensation for bearing this risk is determined by the value of λ_1. The remaining factors in the model are responsible for all of the cross-sectional variation in covariance risk, and hence all of the cross-sectional variation in conditional expected excess stock returns.[3]

2.4. Modeling second-moment dynamics

To complete the model, I need to specify how $\boldsymbol{S}_t$ evolves through time. Equation (2) implies that $\boldsymbol{S}_t$ can be expressed as

$$\boldsymbol{S}_t = \boldsymbol{B}_t \boldsymbol{H}_t \boldsymbol{B}_t' + d_t \boldsymbol{G}_t \boldsymbol{G}_t', \tag{12}$$

so it is sufficient to model the dynamics of $\boldsymbol{H}_t$ and d_t. Because I employ GARCH specifications, I call the resulting process for $\boldsymbol{r}_{t+1}$ a fundamental-factor MGARCH model. The specification for $\boldsymbol{H}_t$ is obtained by decomposing this matrix into conditional factor variances and conditional factor correlations. This facilitates multi-step estimation of the model parameters.

Let $\boldsymbol{h}_t$ denote the $K \times 1$ vector of conditional factor variances for period t (i.e., the main diagonal

[3]Specifying $\boldsymbol{m}_{f,t} = \boldsymbol{\mu}_f$ would yield similar implications with respect to the role of market risk because the conditional loading on the market factor is one for every stock. However, restricting the factor risk premiums to be constant seems less likely to be empirically plausible.

of $\boldsymbol{H}_t$). I assume that this vector evolves as

$$\boldsymbol{h}_t = \boldsymbol{\omega}_h + \boldsymbol{\beta}_h \circ \boldsymbol{h}_{t-1} + \boldsymbol{\alpha}_h \circ \boldsymbol{v}_t^{\circ 2}, \tag{13}$$

where $\boldsymbol{\omega}_h = (\omega_{h,1}, \ldots, \omega_{h,K})'$, $\boldsymbol{\beta}_h = (\beta_{h,1}, \ldots, \beta_{h,K})'$, and $\boldsymbol{\alpha}_h = (\alpha_{h,1}, \ldots, \alpha_{h,K})'$ are $K \times 1$ vectors, $\circ$ denotes the Hadamard (element-by-element) product, and $\boldsymbol{v}_t^{\circ 2} = \boldsymbol{v}_t \circ \boldsymbol{v}_t$ denotes the Hadamard square of $\boldsymbol{v}_t$. In other words, I assume that each factor follows a GARCH(1,1) process that displays GARCH-in-mean effects of the form specified earlier (i.e., $\boldsymbol{v}_{t+1} = \boldsymbol{f}_{t+1} - \boldsymbol{H}_t \boldsymbol{L}_t \boldsymbol{\lambda}$).

Let $\boldsymbol{R}_t = (\boldsymbol{H}_t \circ \boldsymbol{I})^{-1/2} \boldsymbol{H}_t (\boldsymbol{H}_t \circ \boldsymbol{I})^{-1/2}$ denote the conditional correlation matrix of $\boldsymbol{f}_{t+1}$. To facilitate correlation targeting, I assume that $\boldsymbol{R}_t$ is given by a version of the rotated conditional correlation (RCC) specification of Noureldin et al. (2014). Suppose that $\boldsymbol{\Gamma}^{1/2}$ denotes the symmetric square root of $\boldsymbol{\Gamma} = \mathrm{E}(\boldsymbol{R}_t)$.[4] Under the RCC specification, $\boldsymbol{R}_t$ evolves as

$$\boldsymbol{R}_t = (\boldsymbol{Q}_t \circ \boldsymbol{I})^{-1/2} \boldsymbol{Q}_t (\boldsymbol{Q}_t \circ \boldsymbol{I})^{-1/2}, \tag{14}$$

$$\boldsymbol{Q}_t = \boldsymbol{\Gamma}^{1/2} \boldsymbol{P}_t \boldsymbol{\Gamma}^{1/2}, \tag{15}$$

$$\boldsymbol{P}_t = \boldsymbol{I} + (\boldsymbol{\beta}_c \boldsymbol{\beta}_c')^{\circ 1/2} \circ (\boldsymbol{P}_{t-1} - \boldsymbol{I}) + (\boldsymbol{\alpha}_c \boldsymbol{\alpha}_c')^{\circ 1/2} \circ (\boldsymbol{w}_t \boldsymbol{w}_t' - \boldsymbol{I}), \tag{16}$$

where $\boldsymbol{\alpha}_c$ and $\boldsymbol{\beta}_c$ are $K \times 1$ vectors, $(\cdot)^{\circ 1/2}$ denotes the element-wise square root of its argument, and $\boldsymbol{w}_t = \boldsymbol{\Gamma}^{-1/2} (\boldsymbol{H}_{t-1} \circ \boldsymbol{I})^{-1/2} \boldsymbol{v}_t$ satisfies $\mathrm{E}(\boldsymbol{w}_t \boldsymbol{w}_t') = \boldsymbol{I}$.

Equation (14) mirrors the decomposition of $\boldsymbol{R}_t$ used in the dynamic conditional correlation (DCC) model of Engle (2002). It allows the conditional factor correlations to be modeled in terms of the elements of an auxiliary time-varying matrix $\boldsymbol{Q}_t$. Equation (15) defines this auxiliary matrix to be the rotation of another matrix $\boldsymbol{P}_t$ that satisfies $\mathrm{E}(\boldsymbol{P}_t) = \boldsymbol{I}$ by construction. Equation (16) implies that $\boldsymbol{w}_{t+1}$, which is a vector of standardized and rotated excess returns, follows a diagonal version of the multivariate GARCH process of Engle and Kroner (1995), i.e., a diagonal BEKK model. Correlation targeting can be accomplished by substituting a simple moment-based estimator for $\boldsymbol{\Gamma}^{1/2}$ that is consistent under the specified process for the conditional factor variances.

Noureldin et al. (2014) emphasize that the RCC model has two important advantages relative to competing specifications. First, the model guarantees that $\boldsymbol{R}_t$ is positive definite under simple

[4] That is, $\boldsymbol{\Gamma}^{1/2} = \boldsymbol{\Pi} \boldsymbol{\Lambda}^{1/2} \boldsymbol{\Pi}'$, where $\boldsymbol{\Pi} \boldsymbol{\Lambda} \boldsymbol{\Pi}'$ is the eigendecomposition of $\mathrm{E}(\boldsymbol{R}_t)$.

parameter restrictions that can be easily enforced during estimation. It is sufficient to ensure that the inequality restriction $\boldsymbol{\alpha}_c + \boldsymbol{\beta}_c < \mathbf{1}$ is satisfied. Second, parsimonious parameterizations of the $\boldsymbol{P}_t$ process translate into rather rich dynamics for the conditional correlation matrix. The diagonal parameterization in equation (16), for example, implies that $\boldsymbol{Q}_t$ follows a full version of the BEKK model, i.e., a version that has dense asymmetric parameter matrices.

To motivate the specification for d_t, note that the fundamental factor structure implies that $\mathrm{E}(\boldsymbol{e}_{t+1}'\boldsymbol{e}_{t+1}|\mathcal{I}_t) = (N-K)d_t$. Thus we can view d_t as the conditional expected value of the cross-sectional average of squared errors for period $t+1$. If we assume that the cross-sectional average of the squared errors follows an ARMA(1,1) process, then d_t evolves as

$$d_t = \omega_d + \beta_d d_{t-1} + \alpha_d \left(\frac{\boldsymbol{e}_t'\boldsymbol{e}_t}{N-K} \right), \tag{17}$$

where ω_d, β_d, and α_d are scalars. This approach is equivalent to specifying a univariate GARCH(1,1) process for each of the $N-K$ elements of $\boldsymbol{e}_{t+1}$, and then restricting the value of each parameter that appears in the conditional variance equation to be the same for every process.

2.5. Log likelihood calculations

The fundamental factor structure of the model greatly facilitates estimation and inference. Suppose, for example, that the dataset consists of excess returns on N individual stocks for periods $t = 1, 2, \ldots, T$.[5] Let $\boldsymbol{\theta}$ denote the vector of unknown parameters. Under the assumption $\boldsymbol{r}_{t+1}|\mathcal{I}_t \sim N(\boldsymbol{m}_t, \boldsymbol{S}_t)$, the model implies that factors and errors for period $t+1$ are distributed as

$$\left[\begin{array}{c} \boldsymbol{f}_{t+1} \\ \boldsymbol{e}_{t+1} \end{array} \middle| \mathcal{I}_t \right] \sim N\left(\begin{bmatrix} \boldsymbol{m}_{f,t} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \boldsymbol{H}_t & \mathbf{0} \\ \mathbf{0} & d_t \boldsymbol{I} \end{bmatrix} \right), \tag{18}$$

where I have suppressed the dependence of $\boldsymbol{m}_{f,t}$, $\boldsymbol{H}_t$, and d_t on $\boldsymbol{\theta}$ for notational convenience. One can therefore estimate $\boldsymbol{\theta}$ by maximizing

$$\mathcal{L}(\boldsymbol{\theta}) = \sum_{t=1}^{T} -\frac{1}{2}\log|\boldsymbol{H}_{t-1}| - \frac{1}{2}\boldsymbol{v}_t'\boldsymbol{H}_{t-1}^{-1}\boldsymbol{v}_t - \frac{N-K}{2}\log d_{t-1} - \frac{1}{2}\frac{\boldsymbol{e}_t'\boldsymbol{e}_t}{d_{t-1}}, \tag{19}$$

[5] I assume a balanced panel for ease of illustration. The same approach works if the cross-sectional dimension of the panel changes over time.

which is, apart from an additive constant, the log likelihood function for the model. Because $\mathcal{L}(\boldsymbol{\theta})$ is additively separable, the parameters that characterize the dynamics of $\boldsymbol{H}_t$ can be estimated independently of those that characterize the dynamics of d_t with no loss of efficiency.

Although the computational complexity of using numerical methods to maximize $\mathcal{L}(\boldsymbol{\theta})$ obviously depends on the value of N, the rate at which the complexity increases with N is quite slow provided that the value of K is not too large. Indeed, $\mathcal{L}(\boldsymbol{\theta})$ can be computed with little effort beyond that required to fit a sequence of T cross-sectional regressions with sample size N, because a cross-sectional regression of $\boldsymbol{r}_t - \boldsymbol{m}_{t-1}$ on $\boldsymbol{B}_{t-1}$ delivers $\boldsymbol{v}_t$, and $\boldsymbol{e}_t'\boldsymbol{e}_t$ is given by the sum of the squared residuals from this regression. To further reduce the computational demands of the model, I employ a multi-step estimation procedure.

2.6. Estimation and inference

Multi-step estimation procedures are common in the MGARCH literature. Although they entail some sacrifice in efficiency, the gains in terms of computational tractability are typically quite substantial. I view this as a favorable trade-off because drawing inferences about the parameters that characterize the dynamics of $\boldsymbol{H}_t$ and d_t is not the main focus of the analysis. The proposed multi-step procedure is inspired by the methods used for DCC models. However, the presence of GARCH-in-mean effects makes it necessary to employ an iterative approach similar to that used by De Santis and Gerard (1997) to conduct MGARCH-based tests of the conditional CAPM. Convergence usually occurs in a reasonable number of iterations because allowing for time variation in the conditional mean typically has little impact on the estimated dynamics of volatility. Indeed, most MGARCH studies simply assume constant means for this reason.

I begin by fitting a simplified version of the FF-MGARCH model in which $\boldsymbol{m}_{f,t} = \boldsymbol{\mu}_f$ and $\boldsymbol{m}_t = \boldsymbol{B}_t\boldsymbol{\mu}_f$. This yields preliminary estimates of $\{\boldsymbol{H}_t\}_{t=0}^{T-1}$, and $\{d_t\}_{t=0}^{T-1}$. The procedure is as follows (note that $\boldsymbol{f}_t = \boldsymbol{B}_{t-1}^{+}\boldsymbol{r}_t$ and $\boldsymbol{e}_t = \boldsymbol{G}_{t-1}'\boldsymbol{r}_t$ for this version of the model).

1. Construct $\boldsymbol{f}_t = (f_{1,t}, f_{2,t}, \ldots, f_{K,t})'$ for $t = 1, 2, \ldots, T$ and compute $\hat{\boldsymbol{\mu}}_f = (1/T)\sum_{t=1}^{T}\boldsymbol{f}_t$. Estimate $\boldsymbol{\omega}_h$, $\boldsymbol{\beta}_h$, and $\boldsymbol{\alpha}_h$ in equation (13) by fitting a sequence of K univariate GARCH(1,1) models, the kth of which assumes that $f_{k,t}|\mathcal{I}_{t-1} \sim N(\hat{\mu}_{k,f}, h_{kk,t-1})$. Let $\{\hat{\boldsymbol{h}}_t\}_{t=0}^{T-1}$ denote the vector sequence of estimated conditional factor variances.
2. Compute $\hat{z}_{k,t} = (f_{k,t} - \hat{\mu}_{k,f})/\hat{h}_{kk,t-1}^{1/2}$ for $k = 1, 2, \ldots, K$ and $t = 1, 2, \ldots, T$.

3. Compute $\hat{\boldsymbol{\Gamma}} = (1/T)\sum_{t=1}^{T}\hat{\boldsymbol{z}}_t\hat{\boldsymbol{z}}_t'$, the sample covariance matrix of $\hat{\boldsymbol{z}}_t = (\hat{z}_{1,t}, \hat{z}_{2,t}, \ldots, \hat{z}_{K,t})'$. Let $\hat{\boldsymbol{\Gamma}}^{1/2}$ denote the symmetric square root of this matrix.

4. Compute $\hat{\boldsymbol{w}}_t = \hat{\boldsymbol{\Gamma}}^{-1/2}\hat{\boldsymbol{z}}_t$ for $t = 1, 2, \ldots, T$. Estimate $\boldsymbol{\alpha}_c$ and $\boldsymbol{\beta}_c$ in equation (16) by assuming that $\hat{\boldsymbol{z}}_t|\mathcal{I}_{t-1} \sim N(\mathbf{0}, \boldsymbol{R}_{t-1})$. Let $\{\hat{\boldsymbol{R}}_t\}_{t=0}^{T-1}$ denote the estimated sequence of conditional factor correlation matrices.

5. Construct $\{\boldsymbol{e}_t\}_{t=1}^{T}$. Estimate ω_d, β_d, and α_d in equation (17) by assuming that $\boldsymbol{e}_t|\mathcal{I}_{t-1} \sim N(\mathbf{0}, d_{t-1}\boldsymbol{I})$. Let $\{\hat{d}_t\}_{t=0}^{T-1}$ denote the estimated sequence of conditional error variances.

6. Combine $\{\hat{\boldsymbol{h}}_t\}_{t=0}^{T-1}$ with $\{\hat{\boldsymbol{R}}_t\}_{t=0}^{T-1}$ to obtain $\{\hat{\boldsymbol{H}}_t\}_{t=0}^{T-1}$, the estimated sequence of conditional factor covariance matrices.

The time required to complete the procedure is largely governed by step four, which entails optimization over $2K$ parameters and requires T numerical inversions of a $K \times K$ matrix in order to compute the value of the log likelihood function implied by $\hat{\boldsymbol{z}}_t|\mathcal{I}_{t-1} \sim N(\mathbf{0}, \boldsymbol{R}_{t-1})$.

Once the sequence $\{\hat{\boldsymbol{H}}_t\}_{t=0}^{T-1}$ is in hand, a preliminary estimate of $\boldsymbol{\lambda}$ can be obtained by minimizing the criterion

$$\mathcal{Q}(\boldsymbol{\lambda}) = \sum_{t=1}^{T}\hat{\boldsymbol{v}}_t'\hat{\boldsymbol{H}}_{t-1}^{-1}\hat{\boldsymbol{v}}_t, \tag{20}$$

where

$$\hat{\boldsymbol{v}}_t = \boldsymbol{f}_t - \hat{\boldsymbol{H}}_{t-1}\hat{\boldsymbol{L}}_{t-1}\boldsymbol{\lambda} \tag{21}$$

denotes the $K \times 1$ vector of estimated factor innovations obtained by replacing $\boldsymbol{H}_{t-1}$ and $\boldsymbol{L}_{t-1}$ with the preliminary estimates of these matrices. This approach maximizes the log likelihood function in equation (19) for the case in which all the parameters except for $\boldsymbol{\lambda}$ are restricted to equal their preliminary estimates. It is equivalent to pooled generalized-least-squares (GLS) estimation of a sequence of T cross-sectional regressions of the form

$$\boldsymbol{r}_t = \hat{\boldsymbol{C}}_{t-1}^{*}\boldsymbol{\lambda} + \boldsymbol{\varepsilon}_t \tag{22}$$

where $\hat{\boldsymbol{C}}_{t-1}^{*} = \boldsymbol{B}_{t-1}\hat{\boldsymbol{H}}_{t-1}\hat{\boldsymbol{L}}_{t-1}$.[6] To update $\{\hat{\boldsymbol{H}}_t\}_{t=0}^{T-1}$ via a second iteration, I replace $\hat{\mu}_{k,f}$ with the

[6] To see this, consider the cross-sectional regression for period t. The GLS estimator for this regression is obtained by minimizing the criterion

$$\tilde{\mathcal{Q}}_t(\boldsymbol{\lambda}) = \boldsymbol{\varepsilon}_t'\hat{\boldsymbol{S}}_{t-1}^{-1}\boldsymbol{\varepsilon}_t,$$

kth element of $\hat{\boldsymbol{H}}_{t-1}\hat{\boldsymbol{L}}_{t-1}\hat{\boldsymbol{\lambda}}$, where $\hat{\boldsymbol{\lambda}}$ denote the preliminary estimate of $\boldsymbol{\lambda}$. All subsequent updates of $\{\hat{\boldsymbol{H}}_t\}_{t=0}^{T-1}$ are carried out in the same manner using $\hat{\boldsymbol{\lambda}}$ from the previous iteration.

Asymptotic standard errors for the final parameter estimates can be computed by nesting the multi-step estimation procedure within the generalized-method-of-moments framework (see, e.g., Noureldin et al., 2014). But there is a simpler approach that should be adequate for assessing the precision of the parameter estimates that characterize the dynamics of the fitted conditional variances, fitted conditional covariances, and fitted conditional correlations. Specifically, the outer-product and second-derivative estimates of the information matrix for the univariate GARCH(1,1) models in step one and the RCC model in step four can be used to compute quasi maximum likelihood standard errors for the dynamic parameters. The generalized-method-of-moments approach might therefore be reserved for cases in which one wants to perform formal hypothesis tests.

3. Data and Preliminary Analysis

I obtain monthly data for individual stocks from the Center for Research in Security Prices (CRSP) monthly stock file. The sample begins in July 1963, ends in December 2016, and is restricted to ordinary common equity (CRSP share code 10 or 11) for NYSE, AMEX, and NASDAQ firms. The monthly risk-free rate series is taken from the "Fama/French 3 factors" dataset that is posted to the web-based data library maintained by Ken French.[7] The annual data for the required accounting variables are drawn from the Compustat annual industrial file.[8]

I combine the variables into a single dataset by matching accounting information for firms whose fiscal year ends in month t with excess stock returns for months $t+5$ to $t+16$. Under this matching strategy, the accounting variables are lagged by a minimum of four months with respect to the start of the holding period over which the excess stock returns are measured. Four months should be sufficient time for the accounting variables to enter the public information set. This timing convention is also employed in the recent study of Lewellen (2015), which uses long-run averages

where $\hat{\boldsymbol{S}}_{t-1}^{-1} = \boldsymbol{B}_{t-1}^{+\prime}\hat{\boldsymbol{H}}_{t-1}^{-1}\boldsymbol{B}_{t-1}^{+} + \hat{d}_{t-1}^{-1}\boldsymbol{G}_{t-1}\boldsymbol{G}_{t-1}^{\prime}$. By noting that $\boldsymbol{G}_{t-1}\boldsymbol{G}_{t-1}^{\prime}\boldsymbol{B}_{t-1} = \boldsymbol{0}$, it follows that $\tilde{\mathcal{Q}}_t(\kappa, \boldsymbol{\lambda})$ can be expressed as

$$\tilde{\mathcal{Q}}_t(\boldsymbol{\lambda}) = (\boldsymbol{f}_t - \hat{\boldsymbol{H}}_{t-1}\hat{\boldsymbol{L}}_{t-1}\boldsymbol{\lambda})^{\prime}\hat{\boldsymbol{H}}_{t-1}^{-1}(\boldsymbol{f}_t - \hat{\boldsymbol{H}}_{t-1}\hat{\boldsymbol{L}}_{t-1}\boldsymbol{\lambda}) + \hat{d}_{t-1}^{-1}(\boldsymbol{r}_t^{\prime}\boldsymbol{G}_{t-1}\boldsymbol{G}_{t-1}^{\prime}\boldsymbol{r}_t).$$

Hence, minimizing $\sum_{t=1}^{T}\tilde{\mathcal{Q}}_t(\boldsymbol{\lambda})$ produces the same estimates as minimizing $\mathcal{Q}(\boldsymbol{\lambda})$.

[7] See http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

[8] I exclude firms with less than two years of Compustat data to mitigate the well-known biases that arise from the way in which firms are added to the file.

of the coefficient estimates from Fama and MacBeth (1973) regressions to generate out-of-sample forecasts of the cross-section of expected excess stock returns.

The choice of firm characteristics is motivated by prior research. Table 1 described how the characteristics are constructed. The first four variables could be called the "usual suspects" from the recent asset pricing literature: the logarithm of market equity (log ME), the logarithm of the book-to-market equity ratio (log BE/ME), gross profitability scaled by book assets (GP/BA), and the logarithm of the gross growth rate of book assets (log AG). The use of the log ME and log BE/ME variables to capture cross-sectional differences in expected excess stock returns dates to the seminal studies of Fama and French (1992, 1993). But these variables have recently been combined with measures of firm profitability and investment to extend the Fama and French (1993) three-factor model (e.g., Fama and French, 2015; Hou et al., 2015). I refer to the returns on the hedge portfolios associated with these four variables as the size, value, profitability, and investment factors.

The fifth and six variables have been widely studied in the anomalies literature: current accruals scaled by book assets (CA/BA), and the logarithm of the gross two-year growth rate of split-adjusted shares outstanding (log SG). Interest in the accruals anomaly dates to Sloan (1996), who found a strong cross-sectional relation between current accruals and average stock returns. Around the same time, Loughran and Ritter (1995) found that net stock issues are negatively related to subsequent average stock returns. The origins of these anomalies have been the subject of much debate. In contrast, the seventh variable — net sales scaled by market equity (NS/ME) — has received little attention outside of a recent study by Lewellen (2015). His findings suggest that the NS/ME variable captures cross-sectional differences in average stock returns, but the effect seems to be largely confined to firms that have low market equity. I refer to the returns on the hedge portfolios associated with these three variables as the accruals, offerings, and sales factors.

Table 2 reports the sample mean, sample volatility, sample skewness, and selected sample percentiles for the seven firm characteristics.[9] The sample used to compute the statistics in panel A contains all available NYSE, NASDAQ, and AMEX firms (the "full sample"). There is nothing in the results that looks to be cause for concern. In most cases, the statistics are closely aligned with

[9]I winsorize all of the characteristics monthly at the 1st and 99th percentiles of their cross-sectional distribution. Using winsorized or trimmed characteristics is a common approach for limiting the influence of outliers in cross-sectional regression studies that focus on asset pricing issues. See, for example, the recent studies of Novy-Marx (2013) and Ball et al. (2015).

those reported by prior studies that use similar data. The log ME variable, for example, has a mean of 4.67 and volatility 2.19. Ball et al. (2015) report values of 4.55 and 1.97 for this variable using data for a largely-overlapping sample period.

The sample used to compute the statistics in panel B is constructed via a month-by-month screening procedure that excludes microcap stocks (the "no-microcaps sample"). The definition of a microcap stock, which follows Fama and French (2008), is one whose market capitalization for the month is below the 20th percentile of the monthly cross-sectional distribution of market capitalization for NYSE firms. Apart from the rightward shift in the distribution of the log ME variable, the most notable change is an increase in the median excess return from 0.00% to 0.77% per month. In general, dropping microcap stocks from the sample has relatively minor effects on the distributional properties of most characteristics. The exception is the NS/ME variable. The sample percentiles indicate that the right tail of its distribution shortens considerably.

3.1. Fama-MacBeth regressions

Table 3 uses Fama-MacBeth regressions to illustrate the cross-sectional explanatory power of the characteristics for excess stock returns. The results in columns (1) to (6) are for the full sample. Those in the remaining columns are for the no-microcaps sample. In each case, I report two sets of estimates to show the effect of excluding financial firms (SIC codes 6000–6999) from the analysis.[10] The dependent variable for the regressions is the excess percentage stock return, and all specifications include an intercept.

First consider the results for a specification that uses the log ME and log BE/ME variables as regressors. The average estimated slopes are -0.15 and 0.27 with t-statistics of -3.57 and 5.18 for the full sample, or -0.16 and 0.30 with t-statistics of -3.63 and 5.68 if I exclude financial stocks. These results are broadly consistent with those reported by Fama and French (1992) and numerous subsequent studies. Note, however, that the results are sensitive to the sample composition. Using the no-microcaps sample, the average estimated slopes are -0.06 and 0.14 with t-statistics of -1.59 and 2.06, or -0.06 and 0.13 with t-statistics of -1.35 and 1.89 if I exclude financial stocks. Thus the evidence of a cross-sectional relation between the two regressors and expected excess stock returns

[10] The SIC codes used to screen firms are from Compustat (item SICH). If the Compustat code is missing, I replace it with the code from CRSP (item SICCD), if available.

is considerably weaker if we drop microcap stocks from the analysis.

The next specification adds the GP/BA and log AG variables to the set of regressors. The average estimated slopes for these two variables are highly statistically significant, regardless of the sample composition. Using the no-microcaps sample, for example, the average estimated slopes are 0.47 and −0.89 with t-statistics of 4.53 and −6.50, or 0.49 and −0.97 with t-statistics of 4.37 and −6.62 if I exclude financial stocks. Adding the GP/BA and log AG variables as regressors also leads to somewhat stronger evidence (larger absolute t-statistics) that the log ME and log BE/ME variables are related to expected excess stock returns for the no-microcaps sample.

For the final specification, I include the entire set of seven characteristics as regressors. All of the average estimated slopes for this comprehensive specification are statistically significant at the 1% level using the full sample. Excluding financial firms produces some changes in the results, primarily with respect to the accruals variable. However, the changes are relatively minor. The explanatory power of the regressions is clearly weaker for the no-microcaps sample, but with the exception of the accruals and sales variables, the average estimated slopes remain statistically significant. In addition, the average estimated slope on the accruals variable becomes statistically significant if financial firms are excluded from the regressions. I therefore follow Fama and French (1992, 1993), and exclude financial firms for the remainder of the analysis.

3.2. Properties of the fundamental factors

Tables 4 contains descriptive statistics for the market, size, value, profitability, investment, accruals, offerings, and sales factors. Panel A reports the sample mean, sample volatility, sample skewness, and selected sample percentiles for the full and no-microcap samples. Panel B reports the sample correlation matrix of the factors. For any given month, the factor realizations are simply excess returns on well-diversified portfolios of the individual stocks. The portfolio weights sum to one for the market factor, and to zero for each of the remaining characteristic-based factors.[11]

The mean of the market factor is 0.85% per month, or about 10% on an annualized basis. This

[11]Recall that the vector of factor realizations for month t is given by $\boldsymbol{f}_t = (\boldsymbol{B}'_{t-1}\boldsymbol{B}_{t-1})^{-1}\boldsymbol{B}'_{t-1}\boldsymbol{r}_t$. Because the first column of $\boldsymbol{B}_{t-1}$ is a vector of ones, and the elements in each of its remaining columns have a mean of zero and a variance of one (the characteristics are standardized in the cross-sectional dimension), it is easy to see that the first row of $(\boldsymbol{B}'_{t-1}\boldsymbol{B}_{t-1})^{-1}\boldsymbol{B}'_{t-1}$ sums to one, and that each of its remaining rows sums to zero. So the first element of $\boldsymbol{f}_t$ is the excess return on a unit-cost portfolio (the equally-weighted market index) and the remaining elements are excess returns on zero-cost portfolios (characteristic-based hedge portfolios).

is more than double the mean for any other factor in terms of magnitude. However, the market factor is unique in the sense that it entails long positions in all of the individuals stocks. Each of the remaining seven factors is the return on a zero-cost portfolio that has negative weights for roughly half of the stocks and positive weights for the others. So one might anticipate finding that these factors have smaller absolute means (and lower volatilities) than the market.

Notice that the pattern of the mean excess returns is consistent with the results of the Fama-MacBeth regressions in Table 3. Specifically, the mean returns are positive for the market, value, profitability, and sales factors, and negative for the size, investment, accruals, and offerings factors. This pattern is expected because the mean excess portfolio returns are the average slope coefficients that would be obtained by fitting Fama-MacBeth regressions using the standard values of the characteristics as regressors. In other words, the mean excess return for a given factor is an estimate of the marginal effect of a one standard deviation increase in the value of the associated characteristic (loading) on conditional expected excess stock returns.

None of the other distributional properties of the factors stands out as especially noteworthy. Several factors display evidence of mild skewness. The size factor, for example, has an estimated skewness of -1.71. But the return distributions appear to be reasonably symmetric on the whole. The sample correlations between the factors are fairly low in general. The market and offerings factors have the largest correlation at 55%. However, the majority of the correlations are less than 20% in magnitude. The correlations are relatively weak because every factor except one is explicitly designed to capture the marginal explanatory power of a particular characteristic for the cross-section of excess stock returns. In general, the weak correlations are broadly consistent with the view that each factor represents a distinct source of common variation in excess returns.

4. Parameter Estimates and the Cross-Section of Covariance Risk

Table 5 reports estimates of the parameters that determine the dynamics of the conditional variances and conditional covariances under the FF-MGARCH model. I begin with the estimates for the full sample, which are shown in the first six columns of the table. The upper section presents the GARCH(1,1) estimates for the conditional factor variances. As anticipated, the results indicate that the conditional variances are quite persistent. The sum of $\hat{\beta}_{h,k}$ and $\hat{\alpha}_{h,k}$ (the estimated coefficients on the lagged conditional variance and lagged squared demeaned excess return for the kth factor) ranges from 0.74 to 0.97. This finding is in line with evidence from the volatility modeling literature, which overwhelmingly points to strong persistence in stock return volatility.

Note, however, that several of the $\hat{\alpha}_{h,k}$ values are larger than those usually reported in studies that fit GARCH(1,1) models to stock returns. Together with the other estimates, these values are indicative of autocorrelations for the squared demeaned factors that are fairly large. For example, the estimates translate into an estimated first-order autocorrelation of 0.52 in the squared demeaned value factor. Typically, the estimates of GARCH(1,1) parameters for stock returns imply a first-order autocorrelation of around 0.2 or lower for squared demeaned returns. Because the factors are excess returns on portfolios with characteristic-based weights, the results indirectly suggest that the characteristics are cross-sectionally related to the volatility of individual excess stock returns.

The parameter estimates for the no-microcaps sample display the same general patterns, but some changes are evident. First, the estimates of volatility persistence display less variation across factors. The sum of $\hat{\beta}_{h,k}$ and $\hat{\alpha}_{h,k}$ ranges from 0.90 to 0.98. Second, the estimates of $\hat{\alpha}_{h,k}$ tend to be smaller than those obtained using the full sample. But the results still point to autocorrelations for the squared demeaned factors that are fairly large. In the case of the value factor, for instance, the estimates translate into an estimated first-order autocorrelation of 0.53. Overall the evidence suggests that the volatility dynamics of the factors extracted from the no-microcaps sample are similar to those of the factors extracted from the full sample.

The middle section of Table 5 presents the GARCH estimates for conditional error variances. The sum of $\hat{\beta}_d$ and $\hat{\alpha}_d$ for the full sample is 0.97, which is again consistent with strong volatility persistence. Although the estimated value of $\hat{\alpha}_d$, which is 0.42, may seem unusually large, the interpretation of this estimate differs from that of $\hat{\alpha}_h$ in the GARCH(1,1) specifications. Recall that α_d determines how d_t responds to an increase in $\boldsymbol{e}_t'\boldsymbol{e}_t/(N-K)$. Under the fundamental factor structure, $\boldsymbol{e}_t'\boldsymbol{e}_t/(N-K)$ is akin to a realized variance that converges in probability to d_{t-1} as $N \to \infty$. Hence, it is not surprising to find that the weight placed on this quantity is relatively large. Dropping microcap stocks from the sample has little impact on either of the estimates.

The lower section of Table 5 presents the estimates for the RCC specification of conditional factor correlation matrix. The results suggest that the conditional factor correlations display a level of persistence on par with that of the conditional factors variances. In particular, the elements of $(\hat{\boldsymbol{\alpha}}_c\hat{\boldsymbol{\alpha}}_c')^{1/2} + (\hat{\boldsymbol{\beta}}_c\hat{\boldsymbol{\beta}}_c')^{1/2}$, which determine the estimated persistence of the conditional covariances between the rotated versions of the standardized factors, range from 0.57 to 1.00 (rounded to two decimal places), with most exceeding 0.8. In addition, most of the estimated persistence measures increase when microcap stocks are excluded from the sample.

4.1. Cross-section of covariance risk

In view of the parameter estimates, it seems likely that the elements of $\boldsymbol{H}_t$ experience long swings away from their unconditional expected values. The FF-MGARCH model implies that, in general, such swings are accompanied by a shift in the entire cross-section of conditional covariance risk for individual stocks. But how well do the firm characteristics capture cross-sectional differences in covariance risk? I use density plots to provide some initial insights on this question. First, I randomly pick 250 stocks each month from the set of available stocks for the month. Second, I record the estimated conditional covariances between the excess returns on the selected stocks and the fundamental factors for every month in the sample period. Third, I estimate the unconditional density of conditional covariances with each factor. Figure 1 plots these density estimates for both the full and no-microcaps samples.

The plot for the market factor has some interesting implications. Under the FF-MGARCH model, all of the cross-sectional variation in the conditional covariance of excess stock returns with the market factor is explained by the conditional correlations between this factor and the characteristic-based factors (the market factor has a loading of 1 for every stock). Yet the conditional covariance of excess returns with the market displays wide variation across stocks. Indeed, it displays more cross-sectional variation than the conditional covariance of excess returns with any other factor. Dropping microcap stocks from the sample shifts the plot to the left, but the magnitude of the shift is relatively small. Thus the model points to substantial cross-sectional variation in conditional CAPM betas that is directly tied to cross-sectional differences in the values of the firm characteristics. This implication of the model is examined in more depth later on.

Because the conditional loadings for the remaining factors are standardized, the density plots can be used to get a rough idea of the relative importance of each factor in capturing the comovements of excess stock returns. The size and value factors appear to display stronger explanatory power than the other factors, although much of the strength of the size factor appears to be associated with microcap stocks. But the plots suggest that all of the factors have some ability to capture the comovements of excess returns. Of course these visual comparisons are only suggestive.

To dig deeper, I compare the model-based estimates of the *unconditional* covariances to the corresponding sample covariances. Consider, for example, the unconditional covariance with the market factor. I start by forming 50 equally-weighted portfolios that are rebalanced every month. The portfolios are constructed by sorting the set of available stocks for the month in ascending order of the estimated conditional covariance with the market. Once I have determined the composition of each portfolio for every month in the sample period, I use the fitted conditional moments for individual excess stock returns to construct the fitted conditional expected excess portfolio returns and the fitted conditional covariances of the excess portfolio returns with the market factor. Finally, I use the fitted conditional moments of the excess portfolio returns to derive estimates of the unconditional covariances of the excess portfolio returns with the market factor.[12]

Figure 2 illustrates the relation between the model-based estimates of the unconditional covariances and the corresponding sample covariances. If the FF-MGARCH model is correctly specified, then the observed differences between the two sets of covariance estimates arise solely from estimation error. For the most part, the model-based estimates of the unconditional covariances line up pretty well with the sample covariances. The most notable exception is for the size factor. The model-based estimates for this factor are larger than the sample covariances at both ends of the spectrum. This finding suggests a moderate degree of specification error with respect to the size factor, perhaps due to the presence of neglected nonlinearity in the relation between log ME and covariance risk. On the whole, however, the comparisons are not unfavorable to the model.

It is also clear that the extent to which the sorting procedure spreads the unconditional covariances depends very much on the factor under consideration. Sorting on the estimated conditional covariance with the market produces a wide spread in the estimated unconditional covariances with this factor, which is consistent with the evidence from Figure 1. The range of estimated covariances produced by the size- and value-based sorts is also fairly large. In comparison, the accruals-based sort produces a very modest spread in estimated covariances with the accruals factor.

Although the scatterplots in Figure 2 confirm that the FF-MGARCH model captures cross-sectional differences in covariances risk, they do not tell us whether the differences in covariances

[12]This is straightforward using the relation $\text{cov}(x, y) = E[\text{cov}(x, y|z)] + \text{cov}(\text{E(x|z)}, \text{E(y|z)})$, where x , y, and z are random variables defined on the same probability space.

across firms translate into cross-sectional variation in expected excess stock returns. Even if one presumes that covariance risk is priced, the range of estimated covariances for a given factor may not be a reliable indicator of its importance in explaining the cross-section of expected excess returns, because the associated price-of-risk might be relatively large or quite small. Accordingly, I now turn to an examination of the price-of-risk estimates produced by the model.

5. Covariance Risk and the Cross-Section of Expected Excess Returns

Table 6 reports the price-of-risk estimates for the fundamental factors. All of the estimates for the full sample (panel A) are statistically significant at the 1% significance level, with t-statistics that range from 2.93 to 8.07 in absolute value. Thus the analysis indicates that all of the conditional factor covariances help to explain the cross-sectional variation in excess individual stock excess returns. The magnitude of the estimates ranges from 0.08 for the market factor to 0.31 for the accruals factor, with negative signs for the size, investment, accruals, and offerings factors. As anticipated, the sign of each estimate matches that of the corresponding factor mean in Table 4.

Recall that under the proposed version of exact factor pricing, market risk makes the same contribution to the conditional expected excess return of every stock. The estimated value of this contribution is 0.08 times the estimated conditional variance of the component of the market factor that is conditionally uncorrelated with the remaining factors. To help put this in context, the R-squared for a time-series regression of the excess market return on the remaining factors is nearly 50%. Using this information along with the sample volatility of the excess market return (about 6% per month) suggests that, on average, market risk contributes around $0.08 * 36/2 = 1.4$ percentage points per month to expected excess stock returns. Hence the results point to a substantial reward for bearing the component of market risk that is pervasive across all stocks.

Interestingly, however, the estimates imply that the reward for bearing other types of risk is even higher on a per-unit basis. The estimated price of risk for the accruals factor, for example, is almost quadruple that for the market factor. Although this finding is not necessarily inconsistent with rational pricing, it is difficult to understand why investors would demand to be compensated so highly for the covariance of excess stock returns with an accruals-based factor. There may ultimately be a way to reconcile this finding with the predictions of asset pricing theory, but doing so would require a mechanism that links accruals to systematic risk in a persuasive fashion.

To check the stability of the results, I compare the prices-of-risk estimates obtained using three equal-length subperiods: July 1963 to April 1981, May 1981 to February 1999, and March 1999 to December 2016. The estimates clearly display nonnegligible differences across subperiods. Nonetheless, the basic message of the results is not overly sensitive to the choice of subperiod. The sign of the subperiod estimates is consistent for seven of the eight factors. The lone exception is the sales factor, whose price-of-risk estimate is negative for the first subperiod, and positive for the second and third subperiods. But the estimate for the first subperiod is statistically indistinguishable from zero. The most recent of the three subperiods generally produces the smallest absolute t-statistic for each factor. Nonetheless, I find that six of the eight estimates for this subperiod are statistically significant at the 10% level, and five of these are statistically significant at the 5% level.

The price-of-risk estimates for the no-microcaps sample are generally smaller in magnitude than those for the full sample. This finding suggests that the cross-sectional relation between the characteristics and covariance risk is weaker for stocks that have relatively high market capitalization. Even so, the estimates remain statistically significant for all but two of the factors. The lack of significant results for the value and sales factors may to some extent reflect an overall reduction in the precision of the price-of-risk estimates that is caused by excluding microcap stock from the analysis. Dropping these stocks reduces the number of available observations by around 60%.

5.1. Long-term predictive ability of fitted conditional expected excess returns

As previously noted, it is questionable whether rational pricing can be regarded as a plausible explanation for the results reported in Table 6. Ball et al. (2015) suggest a straightforward way to develop additional insights in this regard. In particular, they point out that long-horizon regressions can assist in differentiating between rational and irrational explanations for the cross-sectional predictive ability of firm characteristics, such as operating profitability, for individual stock returns. Their basic argument is that "mispricing is more likely to be corrected over longer horizons," whereas expected stock returns "are likely to be more stationary and, hence, the informativeness of past profitability measures for future returns is likely to persist longer." Because their long-horizon regressions reveal that the cross-sectional relation between operating profitability and stock returns persists for a large number of years, they conclude that the evidence is "difficult to reconcile with market mispricing being the explanation for operating profitability's predictive power."

Table 7 presents a similar analysis for the FF-MGARCH model. First, I use the model to construct the fitted expected excess return for each stock that appears in the dataset for month $t-l$, where $l \in \{1, 6, 12, 34, 36, 60, 120\}$. Second, I identify the subset of stocks that appear in the dataset for every month between $t-l$ and t. Third, I use the data for this subset of stocks to fit cross-section regressions. The dependent variable is either the excess stock return for month t, or the average excess stock return for months $t-l+1$ to t. The explanatory variables are a constant and the fitted conditional expected excess stock return for month $t-l$. I estimate the regression for every month in which it is feasible, and report the average estimated intercept and average estimated slope for both the full sample (panel A) and the no-microcaps sample (panel B).

The results in columns (1) to (5) of panel A are for the regressions that use monthly excess stock returns as the dependent variable. Column (1) reports the results for one-step-ahead forecasts ($l = 1$). The average estimated intercept is -0.29 with a t-statistic of -1.13, and the average estimated slope is 0.98 with a t-statistic of 9.20. Because the former is statistically indistinguishable from zero and the latter is statistically indistinguishable from one at conventional significance levels, I cannot reject the hypothesis that the one-step-ahead forecasts of excess returns are unbiased and efficient. Although the average value of the regression R^2 is very low (only 1%), this is the anticipated finding in view of previous results. Specifically, the average R^2 produced by the Fama-MacBeth regression that uses the full set of characteristics as regressors, which is shown in column (3) of Table 3, is only 3.4%. It is not surprising, therefore, that the fitted conditional expected excess returns explain only a small fraction of the cross-sectional variation in monthly excess returns.

The average estimated slope and average regression R^2 steadily decrease as the forecast horizon increases from one month to three years. The decline in the average estimated slope is consistent with cross-sectional mean reversion in conditional expected excess stock returns. It indicates that the multi-step-ahead forecasts of excess returns (the fitted values for the regression) display a lower cross-sectional dispersion than the fitted conditional expected excess returns. One expects the average regression R^2 to fall as l increases, but it is notable that the explanatory power of the model, as measured by the t-statistic of the average estimated slope, remains highly statistically significant for all forecast horizons. Thus the cross-sectional relation between the fitted conditional expected excess returns and future realized excess returns persists for at least 3 years.

The results for the regressions that use average excess returns as the dependent variable are shown in columns (6) to (10). Column (6) reports the results for one-step-ahead forecasts of average annual excess returns ($k = l = 12$). The average estimated intercept is -0.12 with a t-statistic of -0.43, and the average estimated slope is 0.93 with a t-statistic of 8.99. The most notable change from results in column (1) is that the regression R^2 increases to 3.1%. This increase is qualitatively consistent with strong persistence in conditional expected excess returns. By averaging the excess stock returns over time we reduce noise, and the reduction in noise should improve the signal-to-noise ratio if the conditional expected excess returns are sufficiently persistent.

The average estimated slope steadily decreases as the horizon used to compute average excess returns increases, which is again suggestive of cross-sectional mean reversion in conditional expected excess stock returns. But the regression R^2 displays the opposite behavior as k increases, steadily rising from 3.1% for one-year average returns to 7.0% for ten-year average returns. If one looks favorably on the argument that mispricing is more likely to be corrected over longer horizons, then the regression evidence runs counter to typical mispricing stories. It is difficult to envision how overreaction or underreaction could generate an increase in the regression R^2 as the horizon used to compute average excess returns increases all the way out to 10 years.

Dropping microcap stocks from the sample used to fit the regressions alters the findings to some extent, but the general message of the results remains the same. The average estimated slope and average regression R^2 decrease as the forecast horizon increases, which is the anticipated finding if there is cross-sectional mean reversion in conditional expected excess stock returns. Conversely, the average regression R^2 increases with the horizon that is used to compute average excess returns. Thus the key takeaways are insensitive to the sample composition.

5.2. Properties of portfolios formed on fitted conditional estimated expected excess returns

Earlier I used portfolio sorts to investigate the implications of the FF-MGARCH model with respect to covariance risk. A similar approach is useful for investigating the model's pricing performance. If the constant-price-off-risk specification captures the cross-sectional variation in expected excess stock returns, then using the fitted conditional expected excess returns to group stocks into portfolios should be an effective strategy for spreading average excess portfolio returns. Suppose, for example, that we want to form 25 portfolios in month t that will be held until month $t + 1$. We can use the fitted value of $\boldsymbol{m}_t$ (the one-step-ahead forecast of $\boldsymbol{r}_{t+1}$ under the model) to sort the available stocks in ascending order of estimated expected excess returns and group them accordingly (bottom 4% in portfolio 1, next 4% in portfolio 2, etc.). By repeating the sorting and grouping process for each month in the sample period, we obtain the desired time series of excess portfolio returns. Table 8 examines the properties of portfolios formed on fitted conditional expected excess returns.

The initial ten columns report the sample mean and sample volatility of the excess portfolio returns, followed by their sample covariances with the market, size, value, profitability, investment, accruals, offerings, and sales factors. The results in the first column highlight the economic significance of the cross-sectional explanatory power of the constant-price-off-risk specification. Despite the low R-squared values reported in Table 7, the sorting scheme produces a wide spread in average excess portfolio returns. The average excess return ranges from a low of -0.35% per month for portfolio 1 to 2.99% per month for portfolio 25. Although the increase in the average excess return with the portfolio number is not quite monotonic, the results leave little doubt that the estimates of $\boldsymbol{m}_t$ convey considerable information about the relative performance of stocks in period $t+1$.

The results also highlight the exceedingly poor fit of the CAPM. The sample covariance between the excess portfolio returns and the market factor is largely flat, showing only slight variation across most of the portfolios. It increases somewhat for portfolios at the upper end of the cross-sectional distribution of average excess returns. But the same is true for portfolios at the lower end of the distribution. The lack of any clear relation between the estimated market betas and average excess portfolio returns brings the pricing performance of the CAPM into stark focus.

In contrast to the estimated market betas, the sample covariances of several of the characteristic-based factors with the excess portfolio returns display clear cross-sectional trends. For example, the sample covariances rise steadily as the portfolio number increases for the value and sales factors, and fall steadily for the investment factor. In other cases there appears to be a non-monotonic relation between the sample covariances and portfolio number. Note, however, that it is difficult to interpret these findings in isolation because the characteristic-based factors are correlated with one another. Altering the covariance risk of a portfolio with respect to a given factor will typically alter its covariance risk with respect to all factors. Thus it is not yet clear whether the observed patterns in the sample covariances are in line with the predictions of the FF-MGARCH model.

The last ten columns of the table report the estimates needed to assess the evidence in this regard. Specifically, they contain the estimates of the unconditional expected excess portfolio returns, the unconditional volatilities of the excess portfolio returns, and the unconditional covariances between the excess portfolio returns and factors that are implied by the fitted conditional means, fitted conditional variances, and fitted conditional covarainces from the FF-MGARCH model. If the FF-MGARCH model is correctly specified, then the fitted conditional moments can be used to construct consistent estimators of the unconditional moments provided that the parameters of the model are estimated consistently. Hence, the differences between the sample moments and the model-based estimates can be used as specification diagnostics.

In general, the results paint a reasonably favorable picture of the pricing performance of the model. Like the average excess portfolio returns, the average values of the fitted conditional expected excess returns increase steadily as the portfolio number increases, rising from -0.91% per month for portfolio 1 to 2.64% per month for portfolio 25. The absolute pricing error is below 20 basis points per month for a majority of portfolios, but it increases somewhat at both ends of the cross-sectional distribution of average excess returns. The largest absolute pricing errors are for the first two portfolios: 56 and 47 basis points per month, respectively.

Similarly, the discrepancies between the sample covariances and the model-based estimates of the unconditional covariances are fairly small in general. Consider, for instance, the observed pattern for the market factor. The sample covariance with the market factor starts at 43.0 for portfolio 1, falls to 35.2 for portfolio 10, and then slowly rises to 49.1 for portfolio 25. In comparison, the model-based estimate of the unconditional covariance starts at 46.4 for portfolio 1, falls to 37.3 for portfolio 8, and then slowly rises to 44.9 for portfolio 25. The results for the other seven factors are similar. It is apparent that, on the whole, the cross-sectional patterns in the model-based estimates mirror those in the sample covariances reasonably well.

Figure 3 provides additional evidence on the pricing performance of the model. It illustrates how the relation between the average excess returns and average fitted conditional expected excess returns changes as the number of portfolios formed via the sorting scheme increases. Some deterioration is apparent as the number of portfolios rises from 25 in the top-left panel to 200 in the bottom-right panel, but the drop off in performance is relatively slow. The sample correlation of the average excess portfolio return with the average value of the fitted conditional expected excess portfolio return falls from 0.943 for 25 portfolios to 0.895 for 200 portfolios. Overall the subset of stocks that have the highest estimated conditional expected excess returns appear to be the most troublesome from a pricing perspective. The average excess return on the portfolio that contains these stocks is substantially higher than the corresponding model-based estimate in every case, and the magnitude of the pricing error becomes larger as the number of portfolios increases.

5.3. Properties of portfolios formed on fitted conditional covariances

One drawback of forming portfolios on the basis of fitted conditional expected excess stock returns is that both the low- and high-numbered portfolios tend to contain stocks that are subject to large estimation errors. To see why, consider a scenario in which every stock has exactly the same expected excess return. The sorting is performed purely on estimation error in this case, so it follows that the error in estimating the expected excess portfolio return is maximized for the bottom and top portfolios. This "error maximization" feature of the sorting scheme is undesirable because it is likely to produce large pricing errors for the bottom and top portfolios that are not an accurate reflection of the true performance of the model.

I therefore consider a second set of 25 portfolios that are formed by sorting on the fitted conditional covariances with the market factor. The motivation for this alternative sorting scheme is simple. Under the FF-MGARCH model, the conditional covariances with the market (or, equivalently, the conditional market betas) have no direct bearing on conditional expected excess stock returns. To the extent that a relation exists, it arises indirectly due to correlations between the market factor and the other factors in the model. Hence, sorting on the fitted conditional covariances should distribute the estimation error more evenly than sorting on fitted conditional expected excess returns. Table 9 summarizes the properties of the resulting portfolios.

The general pattern of the results gives the initial impression of lending support to the CAPM. First, the sample covariance between the excess portfolio returns and the market factor increases monotonically from a low of 20.1 for portfolio 1 to a high of 57.3 for portfolio 25, which translates into the estimated market beta rising from 0.5 to 1.5 based on the estimated volatility of this factor (Table 4). Second, the average portfolio excess return increases from a low of 0.58% per month for portfolio 1 to a high of 1.25% per month for portfolio 25. The increase is not quite monotonic, but it lines up with the estimated market betas pretty well.

However, it is apparent that most — if not all — of the cross-sectional variation in the estimated market betas is explained by firm characteristics. Firms with low estimated market betas tend to have high market capitalizations, low book-to-market ratios, high profitability, low asset growth, low accruals, low share growth, and low sales-to-market ratios, while those with high estimated market betas tend to have the opposite characteristics. The estimated volatility of the excess portfolio returns displays the same general pattern, increasing from a low of 4.15% per month for portfolio 1 to a high of 10.43% per month for portfolio 25. So sorting on the estimated conditional betas is clearly effective from the standpoint of spreading portfolio volatility.

The estimates of the unconditional moments implied by the fitted conditional means, conditional variances, and conditional covariances provide further evidence that the firm characteristics capture the cross-sectional variation in the estimated market betas. The model-based estimate of the unconditional covariance between the excess portfolio returns and the market factor increases monotonically from 24.2 for portfolio 1 to 62.5 for portfolio 25. This pattern mimics that displayed by the sample covariances with the market factor. The correspondence is somewhat looser for the model-based estimates of unconditional expected excess returns. The average value of the fitted conditional expected excess return undergoes a fairly sharp decline from portfolio 20 to portfolio 25, while the average excess portfolio return shows no such pattern. On the whole, however, the model has a good deal of success in replicating the cross-sectional patterns in both the sample covariances with the factors and the average excess portfolio returns.

5.4. Fundamental factors in the APT framework

In view of the evidence from the portfolio sorts, one might wonder how the pricing performance of the FF-MGARCH model compares to that of other models that feature prominently in the asset pricing literature. For instance, the Fama and French (2015) five-factor model, which is an unconditional specification of the form envisioned by the APT of Ross (1976), has recently attracted a lot of attention. It is a natural benchmark in the present setting because the factors consist of the excess value-weighted market return along with the excess returns on four characteristic-based hedge portfolios. Tests of the model typically focus on the statistical significance of the estimated intercepts obtained by regressing excess stock or portfolio returns on the five factors.

To benchmark the performance of the FF-MGARCH model against that of the five-factor model, I regress excess portfolios returns on the fundamental factors, and compare the resulting intercepts to those obtained by fitting analogous regressions for the Fama and French (2015) factors. Table

10 presents the side-by-side comparison. I report results for both the 25 portfolios formed on fitted conditional expected excess stock returns (Table 8) and the 25 portfolios formed on fitted conditional covariances of excess stock returns with the market factor (Table 9). The table omits the estimated slope coefficients and associated t-statistics to conserve space. The estimated intercept for each regression is shown first, followed by its t-statistic and the regression R^2.

Columns (1) to (6) present the results for the 25 portfolios formed on fitted conditional expected excess returns. The specifications that use the fundamental factors as regressors produce a statistically-significant intercept for most of the portfolios. Thus there is ample evidence against exact factor pricing in the APT framework. In addition, the estimated intercepts display an easily discernible pattern across portfolios. Portfolios that have either relatively low or quite high average excess returns produce positive estimated intercepts, while those that have intermediate average excess returns produce negative estimated intercepts.

The results obtained using the Fama and French (2015) factors help to put these findings in perspective. The estimated intercepts for these factors generally have smaller absolute t-statistics that those for the fundamental factors. Indeed, the estimated intercepts are statistically insignificant at the 10% level for all but eight of the portfolios. But a reduction in the statistical significance of the estimated intercepts is not necessarily an indication of an improvement in pricing performance. Note in particular that the R^2 value using the Fama and French (2015) factors is always lower than that obtained with the fundamental factors. Consequently, the standard errors of the estimated intercepts tend to be substantially larger than those obtained with the fundamental factors.

The picture that emerges if we assess goodness of fit using the mean absolute pricing error (MAPE) is considerably more favorable to the FF-MGARCH model. The five-factor model produces a MAPE of 29.3 basis points per month, which is larger than the 21.7 basis points per month obtained with the fundamental factors. In addition, the average pricing error for the five-factor model is 16.6 basis points per month, while that for the fundamental factor specification is less than 0.1 basis points per month. Thus the fundamental factors fare reasonably well in explaining the cross-section of average excess stock returns in a relative sense.

A potential concern with such comparisons is that the portfolios are formed on the fitted conditional expected excess returns produced by the FF-MGARCH model. Perhaps the five-factor model is at an inherent disadvantage in this context because the sorting scheme is designed to maximize the dispersion in average excess returns based on the characteristic-based patterns in the data. This should be less of a concern for the portfolios formed on the estimated conditional covariances of excess returns with the market factor, because the market factor plays no role in explaining cross-sectional differences in expected excess returns under the FF-MGARCH model.

Columns (7) to (12) of Table 10 present the results for this second set of portfolios. It is apparent that the fundamental factors perform better in this case. First, all but three of the estimated intercepts are statistically insignificant at the 10% level. Second, the MAPE is only 8.3 basis points per month. These improvements may seem unremarkable given the reduced spread in the average excess returns relative to the first set of portfolios. But consider the results for the five-factor model. Nine of the estimated intercepts are statistical significant at the 10% level, and the MAPE is 17.8 basis points per month. So we again conclude that the fundamental factors fare reasonably well in explaining the cross-section of average excess stock returns.

6. Conclusions

The three-factor model of Fama and French (1993) has been a mainstay of the empirical asset pricing literature for over 20 years. Building on the idea of using the returns on characteristic-based portfolios as risk factors, I develop a new type of MGARCH model that has a fundamental factor structure for individual excess stock returns. Because the model assumes that the loadings on the fundamental factors are observable firm characteristics, it can easily be estimated for systems that contain thousands of individual stocks. Although it is first and foremost a risk model, exact factor pricing can be incorporated and tested by adopting a suitable specification for the vector of conditional expected excess stock returns.

My empirical investigation of the model's performance reveals some evidence of misspecification. But it would be very surprising if this were not the case. One can hardly expect a parsimoniously-parameterized MGARCH model to unerringly capture the dynamics of excess returns for thousands of individual stocks. The question is not whether the proposed model is misspecified, but whether it provides a reasonably accurate description of the process that generates excess stock returns. The evidence is quite encouraging from this perspective. Accordingly, the FF-MGARCH model represents a promising addition to the small set of existing MGARCH models that are designed to capture time-series changes in conditional variances, covariances, and correlations in high-dimensional settings.

In view of its tractability and demonstrated performance, it is likely to find a host of applications in empirical asset pricing, portfolio selection, risk management, and related areas.

References

Alexander C. (2001), Market Models. John Wiley: New York.

Ball, R., J. Gerakos, J., Linnainmaa, and V. Nikolaev (2015), Deflating profitability, *Journal of Financial Economics* 117, 225–248.

Daniel, K., and S. Titman (1997), Evidence on the characteristics of cross sectional variation in stock returns, *Journal of Finance* 52, 1–33.

Daniel, K., and S. Titman (1998), Characteristics or covariances?, *Journal of Portfolio Management* 24, 24–33.

De Santis, G., and B. Gerard (1997), International asset pricing and portfolio diversification with time-varying risk, *Journal of Finance* 52, 1881–1912.

Engle, R. (1987), Multivariate ARCH with factor structures: Cointegration in variance, University of California, San Diego working paper.

Engle, R. (2002), Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models, *Journal of Business and Economic Statistics*, 20, 339-350.

Engle, R., and K. Kroner (1995), Multivariate simultaneous generalized ARCH, *Econometric Theory* 11, 122–150.

Engle, R., and B. Kelly (2012), Dynamic equicorrelation, *Journal of Business and Economic Statistics*, 30, 212–228.

Fama, E., and K. French (1989), Business Conditions and Expected Returns on Stocks and Bonds, *Journal of Financial Economics*, Vol. 25, pp. 23–49.

Fama, E., and K. French (1992), The Cross-Section of Expected Stock Returns, *Journal of Finance*, Vol. 47, pp. 427–465.

Fama, E., and K. French (1993), Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Financial Economics*, Vol. 33, pp. 3–56.

Fama, E., and K. French (1996), Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55–84.

Fama, E., and K. French (2000), Characteristics, covariances, and average returns: 1929 to 1997, *Journal of Finance* 55, 389–406.

Fama, E., and K. French (2008), Average returns, B/M, and share issues. *Journal of Finance* 63, 2971–2995.

Fama, E., French, K., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22.

Fama, E., and J. MacBeth (1973), Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy*, 81, 607–636.

Fan, J., M. Wang, Q. Yao (2008), Modelling multivariate volatilities via conditionally uncorrelated components, *Journal of the Royal Statistical Society, Series B*, 70, 679–702.

Hou, K., C. Xue, and L. Zhang (2015), Digesting anomalies: An investment approach, *The Review of Financial Studies* 28, 650–705.

Lanne, M., and P. Saikkonen (2007), A multivariate generalized orthogonal factor GARCH model, *Journal of Business and Economic Statistics* 25, 61–75.

Lewellen, J. (2015), The cross-section of expected stock returns, *Critical Finance Review* 4, 1–44.

Lintner, J. (1965), The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. Review of Economics and Statistics 47, 13–37.

Loughran, T., and J. Ritter (1995), The new issues puzzle, *Journal of Finance* 50, 23–51.

Markowitz, H. (1956), The optimization of a quadratic function subject to linear constraints. *Naval Research Logistics Quarterly*, 3, 111–133.

Noureldin, D., N. Shephard, and K. Sheppard, (2014), Multivariate rotated ARCH models, *Journal of Econometrics*, 179, 16–30

Novy-Marx, R. (2013), The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1–28.

Sloan, R. (1996), Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71, 289–315.

Rosenberg, B. (1974), Extra-market components of covariance in security returns, *Journal of Financial and Quantitative Analysis*, 9, 263–274.

Ross, S., 1976. The arbitrage theory of capital asset pricing. Journal of Economic Theory 13, 341–360.

Sharpe, W. (1964), Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* 19, 425–442.

Sheather, S., and M. Jones (1991), A reliable data-based bandwidth selection method for kernel density estimation, *Journal of the Royal Statistical Society, Series B* 53, 683–690.

Tipping, M., and C. Bishop (1999), Probabilistic principal component analysis, *Journal of the Royal Statistical Society, Series B* 61, 611–622.

van der Weide, R. (2002), GO-GARCH: A multivariate generalized orthogonal GARCH model, *Journal of Applied Econometrics* 17, 549–564.

Table 1
Firm characteristics used as factor loadings

log ME	The first characteristic is the logarithm of market equity (ME). The ME of a firm is its stock price (CRSP item PRC) multiplied by the number of shares outstanding (CRSP item SHROUT). It is measured in millions of dollars. This characteristic is updated monthly, and is assumed to be known immediately (i.e., the loading on the size factor for month $t+1$ depends on the ME for month t).
log BE/ME	The second characteristic is the logarithm of book equity (BE) to market equity. The BE of a firm is shareholders equity (Compustat item SEQ), plus balance-sheet deferred taxes and investment tax credit (Compustat item TXDITC), if available, minus the book value of preferred stock, which is either its redemption value (Compustat item PSTKRV), liquidation value (Compustat item PSTKL), or par value (Compustat item PSTK), in this order of preference. This definition follows Fama and French (1992). The BE is updated annually, and is assumed to be known four months after the fiscal year end (i.e., the loading on the value factor for month $t+1$ depends on the BE for month $t-4$ or earlier). If the shareholders equity is missing, I substitute common equity plus preferred stock (Compustat item CEQ plus item PSTK), if available, or total assets minus total liabilities (Compustat item AT minus item LT), if available, in this order of preference.
GP/BA	The third characteristic is gross profits (GP) to beginning-of-year book assets (BA). The GP of a firm is the difference between total revenue and cost of goods sold (Compustat item REVT minus item COGS). The BA of a firm is its total assets (Compustat item AT). This characteristic is updated annually, and is assumed to be known four months after the fiscal year end (i.e., the loading on the profitability factor for month $t+1$ depends on the GP for month $t-4$ or earlier).
log AG	The fourth characteristic is the logarithm of the gross asset growth rate over the fiscal year (AG). Assets are book assets (Compustat item AT). This characteristic is updated annually, and is assumed to be known four months after the fiscal year end (i.e., the loading on the investment factor for month $t+1$ depends on the AG for month $t-4$ or earlier).
CA/BA	The fifth characteristic is current accruals (CA) to beginning-of-year book assets. The CA of a firm is the annual change in working capital. Working capital is current assets net of cash (Compustat item ACT minus item CHE) minus current liabilities net of long-term debt (Compustat item LCT minus item DLC). This characteristic is updated annually, and is assumed to be known four months after the fiscal year end (i.e., the loading on the accruals factor for month $t+1$ depends on the CA for month $t-4$ or earlier).
log SG	The sixth characteristic is the logarithm of the gross growth rate of split-adjusted common shares over the prior two fiscal years (SG). Split-adjusted common shares is common shares outstanding times the factor to adjust shares for stock splits (Compustat item CSHO multiplied by item AJEX). This characteristic is updated annually, and is assumed to be known four months after the fiscal year end (i.e., the loading on the offerings factor for month $t+1$ depends on the SG for month $t-4$ or earlier).
NS/ME	The seventh characteristic is net sales (NS) to market equity (NS is Compustat item SALE). The numerator is updated annually, and is assumed to be known four months after the fiscal year end (i.e., the loading on the sales factor for month $t+1$ depends on the NS for month $t-4$ or earlier.

The table describes the CRSP and Compustat data items that define the firm characteristics. The standardized values of the characteristics are used as factor loadings in the FF-MGARCH model. I refer to the returns on the associated hedge portfolios as the size, value, profitability, investment, accruals, offerings, and sales factors, respectively.

Table 2
Descriptive statistics for individual stock returns and firm characteristics

Panel A: All firms

	Mean	Vol	Skew	Percentiles: 1st	10th	25th	50th	75th	90th	99th
Return (%)	1.24	18.01	6.71	−38.60	−15.28	−6.67	0.00	7.29	17.36	57.32
log ME	4.67	2.19	0.30	0.33	1.93	3.06	4.52	6.17	7.63	10.12
log BE/ME	−0.72	1.16	−0.90	−4.02	−2.28	−1.24	−0.52	0.04	0.52	1.44
GP/BA	0.36	0.42	0.70	−0.74	0.02	0.11	0.32	0.56	0.85	1.67
log AG	0.12	0.38	1.91	−0.79	−0.19	−0.02	0.08	0.20	0.47	1.67
CA/BA	−0.11	0.28	−3.37	−0.86	−0.45	−0.15	−0.05	0.00	0.09	0.47
log SG	0.11	0.40	4.58	−0.35	−0.16	−0.01	0.01	0.11	0.39	2.10
NS/ME	2.34	4.01	4.94	0.00	0.13	0.42	1.06	2.52	5.53	20.10

Panel B: No microcaps

	Mean	Vol	Skew	Percentiles: 1st	10th	25th	50th	75th	90th	99th
Return (%)	1.08	12.32	1.21	−31.66	−11.63	−4.97	0.77	6.75	13.93	36.61
log ME	6.62	1.56	0.41	3.68	4.67	5.47	6.53	7.60	8.71	10.72
log BE/ME	−0.84	0.97	−0.96	−3.88	−2.07	−1.32	−0.70	−0.19	0.22	0.92
GP/BA	0.38	0.37	1.15	−0.46	0.05	0.13	0.32	0.55	0.83	1.58
log AG	0.16	0.31	3.03	−0.60	−0.05	0.03	0.10	0.20	0.42	1.46
CA/BA	−0.11	0.26	−4.18	−0.84	−0.44	−0.12	−0.05	−0.01	0.05	0.32
log SG	0.09	0.34	5.66	−0.31	−0.09	−0.01	0.01	0.10	0.29	1.74
NS/ME	1.48	2.13	6.31	0.00	0.18	0.41	0.87	1.74	3.25	9.98

The table reports the sample mean (Mean), sample volatility (Vol), sample skewness (Skew), and selected sample percentiles for the dataset variables. The statistics in panel A are computed using all available firm-month observations (the "full sample"). The statistics in panel B are computed using the subset of firm-month observations obtained by excluding firms whose market equity for the month is less than the 20th percentile of the monthly cross-sectional distribution of market equity for NYSE firms (the "no-microcaps sample"). The characteristics are the logarithm of market equity in millions (log ME), the logarithm of the ratio of book equity to market equity (log BE/ME), the ratio of gross profits to book assets (GP/BA), the logarithm of the growth in book assets over the year (log AG), the ratio of current accruals to book assets (CA/BA), the logarithm of the growth in split-adjusted common shares over two years (log SG), and the ratio of sales to market equity (NS/ME). All accounting variables are updated four months after the end of the firm's fiscal year, and all of the characteristics are winsorized monthly at the 1st and 99th percentiles of their cross-sectional distribution. The sample period is July 1963 to December 2016.

Table 3
Fama-MacBeth regressions

$$r_{n,t} = \rho_1 + \rho_{ME} \log ME_{t-1} + \rho_{BM} \log BE/ME_{t-1} + \rho_{GP}\, GP/BA_{t-1} + \rho_{AG} \log AG_{t-1} + \rho_{CA}\, CA/BA_{t-1} + \rho_{SG} \log SG_{t-1} + \rho_{NS}\, NS/ME_{t-1} + \epsilon_{n,t}$$

	All stocks						No microcaps					
	Including financial			Excluding financial			Including financial			Excluding financial		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\bar{\hat{\rho}}_1$	2.04	1.91	1.70	2.10	1.96	1.74	1.52	1.54	1.45	1.46	1.48	1.43
	(5.93)	(5.65)	(5.19)	(5.74)	(5.32)	(5.03)	(3.92)	(4.18)	(4.05)	(3.51)	(3.71)	(3.75)
$\bar{\hat{\rho}}_{ME}$	−0.15	−0.13	−0.12	−0.16	−0.14	−0.13	−0.06	−0.07	−0.07	−0.06	−0.07	−0.07
	(−3.57)	(−3.13)	(−3.06)	(−3.63)	(−3.28)	(−3.12)	(−1.59)	(−1.86)	(−1.98)	(−1.35)	(−1.68)	(−1.79)
$\bar{\hat{\rho}}_{BM}$	0.27	0.28	0.23	0.30	0.29	0.24	0.14	0.15	0.12	0.13	0.13	0.12
	(5.18)	(5.34)	(4.79)	(5.68)	(5.45)	(4.91)	(2.06)	(2.31)	(1.99)	(1.89)	(1.97)	(1.83)
$\bar{\hat{\rho}}_{GP}$		0.51	0.48		0.49	0.45		0.47	0.52		0.49	0.49
		(6.11)	(5.45)		(5.27)	(4.88)		(4.53)	(4.91)		(4.37)	(4.38)
$\bar{\hat{\rho}}_{AG}$		−1.23	−0.95		−1.21	−0.92		−0.89	−0.70		−0.97	−0.73
		(−11.10)	(−9.27)		(−10.98)	(−9.01)		(−6.50)	(−5.53)		(−6.62)	(−5.46)
$\bar{\hat{\rho}}_{CA}$			−0.47			−0.63			−0.36			−0.49
			(−2.51)			(−4.01)			(−1.54)			(−2.22)
$\bar{\hat{\rho}}_{SG}$			−0.33			−0.32			−0.46			−0.47
			(−3.16)			(−3.05)			(−4.67)			(−4.44)
$\bar{\hat{\rho}}_{NS}$			0.05			0.05			0.01			0.01
			(3.56)			(3.89)			(0.76)			(0.63)
$\bar{R}^2$	0.022	0.028	0.034	0.022	0.028	0.033	0.029	0.042	0.055	0.030	0.044	0.054
$\bar{N}$	4072	4072	4072	3435	3435	3435	1668	1668	1668	1408	1408	1408

The table reports time-series averages of selected statistics for monthly cross-sectional regressions. The dependent variable for the regressions is the monthly percentage excess stock return. The regressors are the logarithm of market equity in millions (log ME), the logarithm of the ratio of book equity to market equity (log BE/ME), the ratio of gross profits to book assets (GP/BA), the logarithm of the growth in book assets over the year (log AG), the ratio of current accruals to book assets (CA/BA), the logarithm of the growth in split-adjusted common shares over two years (log SG), and the ratio of net sales to market equity (NS/ME). All accounting variables are updated four months after the end of the firm's fiscal year, and all of the characteristics are winsorized monthly at the 1st and 99th percentiles of their cross-sectional distribution. I fit the cross-sectional regression for every month from July 1963 to December 2016, and report the average values of the estimated coefficients, R-squared statistic, and number of cross-sectional observations. Fama and MacBeth (1973) t-statistics are shown below the average estimated coefficients in parentheses. The results in columns (1) to (6) are for the full set of NYSE, AMEX, and NASDAQ firms. The results in columns (7) to (12) are for the subset of observations obtained by excluding firms whose market equity for the month is less than the 20th percentile of the monthly cross-sectional distribution of market equity for NYSE firms. In each case, I illustrate the impact of excluding financial firms (those with SIC codes from 6000 to 6999) from the regressions.

Table 4
Descriptive statistics for fundamental factors

Panel A: Selected distributional properties

	Factor portfolios include all stocks								Factor portfolios exclude microcap stocks							
				Percentiles								Percentiles				
	Mean	Vol	Skew	5th	25th	50th	75th	95th	Mean	Vol	Skew	5th	25th	50th	75th	95th
f_{MAR}	0.85	6.18	−0.14	−8.86	−2.73	0.94	4.38	9.90	0.67	5.51	−0.47	−8.15	−2.52	1.08	4.34	8.98
f_{SIZ}	−0.25	2.05	−1.71	−3.58	−1.05	−0.01	0.89	2.31	−0.09	1.17	−0.63	−1.99	−0.78	−0.02	0.64	1.68
f_{VAL}	0.26	1.40	−0.58	−1.61	−0.35	0.23	0.84	2.19	0.11	1.58	−0.47	−1.98	−0.63	0.04	0.82	2.23
f_{PRO}	0.18	0.86	−0.77	−1.30	−0.33	0.24	0.74	1.50	0.18	0.96	−0.21	−1.42	−0.42	0.18	0.76	1.78
f_{INV}	−0.33	0.91	−0.63	−1.65	−0.69	−0.28	0.13	0.83	−0.23	1.02	−1.50	−1.73	−0.62	−0.17	0.26	1.13
f_{ACC}	−0.09	0.54	0.07	−0.93	−0.42	−0.09	0.25	0.73	−0.07	0.66	0.41	−1.01	−0.46	−0.09	0.29	1.00
f_{OFF}	−0.08	0.74	0.91	−1.07	−0.46	−0.12	0.25	1.04	−0.08	0.60	0.76	−1.00	−0.41	−0.09	0.21	0.80
f_{SAL}	0.18	1.07	2.04	−1.04	−0.38	0.06	0.53	1.71	0.03	0.84	0.27	−1.28	−0.39	0.03	0.46	1.30

Panel B: Sample correlations between factors

	Factor portfolios include all stocks								Factor portfolios exclude microcap stocks							
	f_{MAR}	f_{SIZ}	f_{VAL}	f_{PRO}	f_{INV}	f_{ACC}	f_{OFF}	f_{SAL}	f_{MAR}	f_{SIZ}	f_{VAL}	f_{PRO}	f_{INV}	f_{ACC}	f_{OFF}	f_{SAL}
f_{MAR}	1.00	−0.46	−0.35	−0.19	0.21	0.15	0.55	0.31	1.00	−0.60	−0.32	−0.05	0.38	0.14	0.36	0.27
f_{SIZ}	−0.46	1.00	0.42	0.38	−0.10	−0.17	−0.53	−0.19	−0.60	1.00	0.38	0.10	−0.27	−0.21	−0.34	−0.12
f_{VAL}	−0.35	0.42	1.00	0.26	−0.25	−0.19	−0.32	0.18	−0.32	0.38	1.00	0.34	−0.45	−0.20	−0.30	0.21
f_{PRO}	−0.19	0.38	0.26	1.00	−0.16	−0.15	−0.33	−0.12	−0.05	0.10	0.34	1.00	−0.25	−0.18	−0.12	0.06
f_{INV}	0.21	−0.10	−0.25	−0.16	1.00	−0.15	0.21	−0.02	0.38	−0.27	−0.45	−0.25	1.00	−0.03	0.21	−0.04
f_{ACC}	0.15	−0.17	−0.19	−0.15	−0.15	1.00	0.02	−0.12	0.14	−0.21	−0.20	−0.18	−0.03	1.00	0.06	−0.07
f_{OFF}	0.55	−0.53	−0.32	−0.33	0.21	0.02	1.00	0.08	0.36	−0.34	−0.30	−0.12	0.21	0.06	1.00	−0.18
f_{SAL}	0.31	−0.19	0.18	−0.12	−0.02	−0.12	0.08	1.00	0.27	−0.12	0.21	0.06	−0.04	−0.07	−0.18	1.00

The table presents descriptive statistics for the market, size, value, profitability, investment, accruals, offerings, and sales factors. I extract the factors from individual excess stock returns by treating the standardized values of the firm characteristics as observable time-varying factor loadings. Specifically, the vector of factors for month t is given by $\boldsymbol{f}_t = (\boldsymbol{B}'_{t-1}\boldsymbol{B}_{t-1})^{-1}\boldsymbol{B}'_{t-1}\boldsymbol{r}_t$, where $\boldsymbol{B}_{t-1}$ is the factor loading matrix (one row per firm) and $\boldsymbol{r}_t$ is the vector of excess stock returns. The first column of $\boldsymbol{B}_{t-1}$ is a vector of ones, and each of the remaining $K-1$ columns has a mean of zero and a variance of one. Thus $\boldsymbol{f}_t$ can be interpreted as a vector of excess returns for a set well-diversified portfolios that have characteristic-based weights. Panel A reports sample mean, sample volatility, sample skewness, and selected sample percentile for the factors. Panel B reports the sample correlation matrix of the factors. In each case, I report statistics for the factors extracted from the full set of NYSE, AMEX, and NASDAQ firms, and those extracted from the subset of observations obtained by excluding firms whose market equity for the month is less than the 20th percentile of the monthly cross-sectional distribution of market equity for NYSE firms. The sample period is July 1963 to December 2016.

Table 5
Estimates of dynamic parameters for the FF-MGARCH model

	All nonfinancial stocks				No microcaps			
	Estimates		Std. Errors		Estimates		Std. Errors	
Factors	β_h	α_h	β_h	α_h	β_h	α_h	β_h	α_h
h_{MAR}	0.83	0.09	0.050	0.037	0.83	0.09	0.054	0.037
h_{SIZ}	0.46	0.28	0.263	0.089	0.82	0.11	0.048	0.032
h_{VAL}	0.54	0.34	0.097	0.066	0.62	0.30	0.080	0.057
h_{PRO}	0.58	0.20	0.158	0.058	0.77	0.13	0.131	0.050
h_{INV}	0.77	0.19	0.047	0.038	0.77	0.19	0.049	0.037
h_{ACC}	0.88	0.09	0.032	0.022	0.88	0.10	0.034	0.025
h_{OFF}	0.77	0.20	0.068	0.054	0.79	0.15	0.078	0.051
h_{SAL}	0.83	0.14	0.038	0.030	0.84	0.12	0.058	0.039
Spherical errors	β_d	α_d	β_d	α_d	β_d	α_d	β_d	α_d
d	0.55	0.42	0.110	0.100	0.57	0.41	0.044	0.043
Standardized & rotated factors	β_c	α_c	β_c	α_c	β_c	α_c	β_c	α_c
$P_{MAR,MAR}$	0.91	0.09	0.031	0.025	0.93	0.05	0.049	0.029
$P_{SIZ,SIZ}$	0.76	0.06	0.085	0.031	0.88	0.04	0.153	0.033
$P_{VAL,VAL}$	0.86	0.05	0.093	0.031	0.87	0.09	0.054	0.033
$P_{PRO,PRO}$	0.98	0.02	0.019	0.008	0.85	0.04	0.080	0.021
$P_{INV,INV}$	0.70	0.08	0.114	0.036	0.65	0.13	0.082	0.054
$P_{ACC,ACC}$	0.91	0.04	0.052	0.021	0.92	0.05	0.047	0.019
$P_{OFF,OFF}$	0.85	0.04	0.107	0.025	0.88	0.04	0.051	0.016
$P_{SAL,SAL}$	0.50	0.07	0.154	0.042	0.97	0.03	0.050	0.014

The table reports estimates of the dynamic parameters for the FF-MGARCH model. I assume that each element of $\boldsymbol{h}_t$, the vector of conditional variances for the market, size, value, profitability, investment, accruals, offerings, and sales factors, follows a GARCH(1,1) process, and incorporate exact factor pricing via GARCH-in-mean effects. The coefficients on the lagged conditional variances are reported under β_d and the coefficients on the lagged squared factor innovations are reported under α_d. I assume that the conditional covariance matrix of the errors, $d_t\boldsymbol{I}$, is described by a GARCH process in which d_t takes the form implied by specifying an ARMA(1,1) model for the cross-sectional average of the squared errors. The coefficient on the lagged value of d_t is reported under β_d and the coefficient on the lagged cross-sectional average of the squared errors is reported under α_d. I assume that the conditional correlation matrix of the factors, $\boldsymbol{R}_t$, follows a multivariate rotated conditional correlation (RCC) process. The dynamics of $\boldsymbol{R}_t$ are specified in terms of a recurrence relation for an auxiliary matrix $\boldsymbol{P}_t$ that evolves as a function of the vector of standardized and rotated factors, $\boldsymbol{w}_t$. The recurrence relation for diagonal elements of $\boldsymbol{P}_t$ takes the same general form as that for the condition GARCH(1,1) variances. The coefficients on the lagged diagonal elements of $\boldsymbol{P}_t$ are reported under β_c and the coefficients on the lagged squared values of the standardized and rotated factors are reported under α_c. The dynamics of $\boldsymbol{P}_t$ are fully determined by these parameters. I estimate the FF-MGARCH model via an iterative, multi-step, likelihood-based procedure that assumes that individual excess stock returns follow a conditional multivariate normal distribution. First, I obtain preliminary estimates of the dynamic parameters by treating the vector of conditional factor means as constant (no GARCH-in-mean effects). Second, I use the fitted conditional variances and fitted conditional correlations to obtain a preliminary estimate of the price-of-risk vector that characterizes the GARCH-in-mean effects. Third, I use the preliminary estimate of the price-of-risk vector to update the estimates of the dynamic parameters. Fourth, I use the updated estimates of the dynamic parameters to update the estimate of the price-of-risk vector. The iterations continue in this fashion until convergence. I report two sets of parameter estimates: one for the full set of NYSE, AMEX, and NASDAQ firms, and another obtained by excluding firms whose market equity for the month is less than the 20th percentile of the monthly cross-sectional distribution of market equity for NYSE firms. I use the quasi maximum likelihood estimator of the asymptotic covariance matrix to compute standard errors. The sample period is July 1963 to December 2016.

Table 6
Estimates of price-of-risk parameters for the FF-MGARCH model

Panel A: All firms

	λ_{MAR}	λ_{SIZ}	λ_{VAL}	λ_{PRO}	λ_{INV}	λ_{ACC}	λ_{OFF}	λ_{SAL}
	0.08	−0.17	0.11	0.27	−0.28	−0.31	−0.17	0.11
	(8.07)	(−7.25)	(3.67)	(5.45)	(−7.08)	(−4.15)	(−2.96)	(2.93)
				Subperiod results				
Jul 63 to Apr 81	0.09	−0.27	0.07	0.16	−0.23	−0.43	−0.55	−0.13
	(5.56)	(−6.51)	(1.19)	(1.88)	(−2.55)	(−3.62)	(−4.32)	(−1.60)
May 81 to Feb 99	0.09	−0.15	0.20	0.52	−0.71	−0.52	−0.05	0.27
	(5.05)	(−3.93)	(3.47)	(6.80)	(−8.68)	(−3.56)	(−0.37)	(3.75)
Mar 99 to Dec 16	0.07	−0.11	0.09	0.16	−0.19	−0.09	−0.09	0.10
	(3.87)	(−2.93)	(2.33)	(1.97)	(−4.17)	(−0.76)	(−1.18)	(1.91)

Panel B: No microcaps

	λ_{MAR}	λ_{SIZ}	λ_{VAL}	λ_{PRO}	λ_{IV}	λ_{ACC}	λ_{OFF}	λ_{SAL}
	0.07	−0.18	−0.03	0.12	−0.27	−0.21	−0.30	−0.03
	(5.91)	(−4.82)	(−1.07)	(2.76)	(−6.21)	(−3.41)	(−4.28)	(−0.52)
				Subperiod results				
Jul ‘63 to Apr ‘81	0.05	−0.29	0.03	0.06	−0.17	−0.09	−0.64	−0.16
	(2.46)	(−4.45)	(0.58)	(0.82)	(−1.88)	(−1.03)	(−5.01)	(−1.77)
May 81 to Feb 99	0.10	−0.08	−0.10	0.27	−0.45	−0.74	−0.32	0.04
	(4.90)	(−1.20)	(−1.88)	(3.73)	(−5.13)	(−5.83)	(−2.33)	(0.38)
Mar 99 to Dec 16	0.06	−0.17	0.00	0.02	−0.23	−0.06	−0.14	−0.03
	(2.99)	(−2.95)	(0.05)	(0.26)	(−4.05)	(−0.51)	(−1.39)	(−0.41)

The table reports estimates of the price-of-risk vector, $\boldsymbol{\lambda}$, that appears in the conditional mean specification for the FF-MGARCH model. The model assumes that each element of the vector of conditional variances for the market, size, value, profitability, investment, accruals, offerings, and sales factors follows a GARCH(1,1) process. It incorporates exact factor pricing via GARCH-in-mean effects. The price of risk for a factor is the incremental contribution to conditional expected excess stock returns per unit of exposure to covariance risk with the factor. I estimate the FF-MGARCH model via an iterative, multi-step, likelihood-based procedure that assumes that individual excess stock returns follow a conditional multivariate normal distribution. First, I obtain preliminary estimates of the dynamic parameters by treating the vector of conditional factor means as constant (no GARCH-in-mean effects). Second, I use the fitted conditional variances and fitted conditional correlations to obtain a preliminary estimate of the price-of-risk vector that characterizes the GARCH-in-mean effects. Third, I use the preliminary estimate of the price-of-risk vector to update the estimates of the dynamic parameters. Fourth, I use the updated estimates of the dynamic parameters to update the estimate of the price-of-risk vector. The iterations continue in this fashion until convergence. I report two sets of parameter estimates: one for the full set of NYSE, AMEX, and NASDAQ firms, and another obtained by excluding firms whose market equity for the month is less than the 20th percentile of the monthly cross-sectional distribution of market equity for NYSE firms. I use the quasi maximum likelihood estimator of the asymptotic covariance matrix to compute standard errors. The sample period is July 1963 to December 2016.

Table 7
Cross-sectional forecasting performance of estimated expected stock returns

Panel A: All firms

$$\frac{1}{k}\sum_{j=1}^{k} r_{n,t+j-k} = \rho_1 + \rho_m \hat{m}_{n,t-l} + \epsilon_{n,t}$$

	Monthly returns ($k=1$)					Average of monthly returns ($k=l$)				
	$l=1$	$l=6$	$l=12$	$l=24$	$l=36$	$l=12$	$l=24$	$l=36$	$l=60$	$l=120$
$\bar{\hat{\rho}}_1$	−0.29	−0.14	0.02	0.27	0.32	−0.12	0.04	0.15	0.31	0.58
	(−1.13)	(−0.53)	(0.10)	(1.10)	(1.28)	(−0.49)	(0.18)	(0.62)	(1.30)	(2.46)
$\bar{\hat{\rho}}_m$	0.98	0.88	0.84	0.66	0.57	0.93	0.89	0.83	0.69	0.49
	(9.20)	(7.50)	(7.70)	(6.25)	(5.00)	(8.99)	(8.71)	(8.28)	(7.29)	(5.36)
$\bar{R}^2(\%)$	1.0	0.9	1.0	0.8	0.7	3.1	4.4	5.4	6.5	7.0
$\bar{N}$	3435.	3293.	3134.	2846.	2595.	3134.	2846.	2595.	2178.	1465.

Panel B: No microcaps

	Monthly returns ($k=1$)					Average of monthly returns ($k=l$)				
	$l=1$	$l=6$	$l=12$	$l=24$	$l=36$	$l=12$	$l=24$	$l=36$	$l=60$	$l=120$
$\bar{\hat{\rho}}_1$	0.09	0.23	0.14	0.23	0.48	0.17	0.22	0.28	0.32	0.44
	(0.38)	(1.03)	(0.67)	(1.04)	(2.11)	(0.78)	(1.04)	(1.33)	(1.57)	(2.16)
$\bar{\hat{\rho}}_m$	0.79	0.68	0.67	0.65	0.45	0.88	0.82	0.76	0.70	0.49
	(4.91)	(4.23)	(4.05)	(3.97)	(2.76)	(5.87)	(5.66)	(5.41)	(5.35)	(4.08)
$\bar{R}^2(\%)$	1.5	1.4	1.4	1.2	1.0	2.9	3.7	3.9	4.5	5.4
$\bar{N}$	1408.	1283.	1194.	1066.	971.	1194.	1066.	971.	833.	612.

The table reports time-series averages of selected statistics for monthly cross-sectional regressions. The dependent variable for the regressions is either the excess percentage stock return for month t, or the average excess percentage stock return for months $t-l+1$ to t. The explanatory variables are a constant and the fitted conditional expected excess percentage stock return for month $t-l$. I fit the regressions as follows. First, I use the FF-MGARCH parameter estimates to compute the fitted conditional expected excess return for each stock that has data for month $t-l$, where $l \in \{1, 6, 12, 34, 36, 60, 120\}$. Second, I identify the subset of stocks that appear in the dataset for every month between $t-l$ and t. Third, I use the data for this subset of stocks to fit a cross-section regression for month t via OLS. I fit the regression for every month from July 1963 to December 2016 for which it is feasible (the number of months with available data depends on the values of k and l), and report the average values of the estimated coefficients, R-squared statistic, and number of cross-sectional observations. Autocorrelation-robust versions of Fama and MacBeth (1973) t-statistics are shown below the average estimated coefficients in parentheses (the monthly estimates of the coefficients are serially correlated by construction for $k > 1$). The estimates reported in panel A are for the full set of NYSE, AMEX, and NASDAQ firms. The estimates reported in panel B are for the subset of firm-month observations obtained by excluding firms whose market equity for the month is less than the 20th percentile of the monthly cross-sectional distribution of market equity for NYSE firms.

Table 8
Properties of excess returns on 25 portfolios formed by sorting on estimated conditional expected stock returns

	Sample estimates of unconditional moments										FF-MGARCH estimates of unconditional moments									
			Covariances with fundamental factors										Covariances with fundamental factors							
№	Mean	Vol	f_{MAR}	f_{SIZ}	f_{VAL}	f_{PRO}	f_{INV}	f_{ACC}	f_{OFF}	f_{SAL}	Mean	Vol	f_{MAR}	f_{SIZ}	f_{VAL}	f_{PRO}	f_{INV}	f_{ACC}	f_{OFF}	f_{SAL}
1	−0.35	8.09	43.01	−6.17	−6.30	−1.85	3.07	0.62	3.72	0.91	−0.91	9.19	46.35	−4.06	−6.92	−1.47	6.13	0.75	3.74	0.25
2	0.15	7.34	40.76	−5.21	−5.17	−1.36	2.53	0.49	3.23	1.04	−0.32	7.73	41.52	−3.01	−5.68	−0.88	4.28	0.63	2.99	0.40
3	0.11	6.73	37.78	−4.69	−4.70	−1.18	2.37	0.47	2.92	0.88	−0.04	7.08	39.52	−2.74	−4.99	−0.69	3.41	0.58	2.65	0.52
4	0.28	6.44	36.73	−4.43	−4.40	−1.07	1.99	0.55	2.68	0.93	0.16	6.69	38.37	−2.67	−4.46	−0.59	2.85	0.56	2.43	0.63
5	0.35	6.26	36.61	−4.45	−3.95	−0.98	1.84	0.54	2.51	1.09	0.31	6.47	37.83	−2.81	−4.12	−0.54	2.48	0.55	2.31	0.73
6	0.39	6.21	36.19	−4.24	−3.95	−0.81	1.70	0.58	2.44	0.98	0.43	6.34	37.49	−2.96	−3.86	−0.53	2.19	0.54	2.24	0.82
7	0.48	6.00	35.35	−3.87	−3.42	−0.68	1.61	0.47	2.24	1.21	0.52	6.25	37.31	−3.17	−3.63	−0.52	2.00	0.54	2.19	0.90
8	0.53	5.99	35.53	−4.14	−3.28	−0.72	1.41	0.52	2.20	1.31	0.61	6.20	37.27	−3.41	−3.46	−0.54	1.84	0.54	2.17	0.98
9	0.56	6.01	35.78	−4.07	−3.22	−0.68	1.37	0.57	2.16	1.34	0.68	6.18	37.31	−3.68	−3.31	−0.56	1.71	0.54	2.17	1.06
10	0.60	5.90	35.15	−3.89	−2.79	−0.63	1.07	0.56	2.09	1.40	0.75	6.17	37.44	−3.96	−3.20	−0.59	1.61	0.54	2.18	1.14
11	0.69	6.04	36.13	−4.35	−2.80	−0.68	1.10	0.52	2.20	1.61	0.82	6.17	37.50	−4.19	−3.08	−0.59	1.51	0.54	2.17	1.21
12	0.72	6.07	36.43	−4.66	−2.80	−0.77	1.02	0.54	2.19	1.65	0.89	6.20	37.75	−4.52	−2.99	−0.63	1.43	0.55	2.20	1.30
13	0.78	6.08	36.35	−4.68	−2.47	−0.66	0.99	0.53	2.11	1.70	0.95	6.21	37.84	−4.76	−2.88	−0.65	1.35	0.55	2.21	1.38
14	0.78	6.05	36.16	−4.89	−2.60	−0.74	0.87	0.55	2.12	1.87	1.01	6.26	38.14	−5.11	−2.81	−0.69	1.28	0.55	2.24	1.47
15	0.91	6.19	37.04	−5.39	−2.65	−0.86	0.82	0.60	2.23	1.91	1.08	6.32	38.42	−5.45	−2.73	−0.72	1.21	0.56	2.28	1.56
16	0.87	6.11	36.37	−5.18	−2.21	−0.74	0.79	0.53	2.14	1.99	1.14	6.36	38.59	−5.73	−2.64	−0.76	1.14	0.55	2.30	1.65
17	0.89	6.12	36.39	−5.58	−2.11	−0.73	0.82	0.51	2.14	2.18	1.21	6.42	38.85	−6.04	−2.55	−0.78	1.07	0.56	2.32	1.75
18	1.04	6.32	37.37	−6.06	−2.17	−0.97	0.67	0.55	2.25	2.32	1.28	6.49	39.16	−6.40	−2.47	−0.82	1.00	0.56	2.36	1.86
19	1.04	6.32	37.37	−6.42	−2.32	−0.91	0.69	0.58	2.28	2.38	1.36	6.58	39.49	−6.76	−2.37	−0.85	0.93	0.55	2.40	2.00
20	1.06	6.38	37.47	−6.75	−2.30	−0.97	0.56	0.55	2.31	2.45	1.44	6.69	39.85	−7.15	−2.26	−0.89	0.85	0.55	2.44	2.16
21	1.22	6.69	39.10	−7.46	−2.19	−1.17	0.64	0.49	2.52	2.76	1.54	6.81	40.25	−7.55	−2.14	−0.92	0.76	0.54	2.49	2.35
22	1.37	6.95	40.16	−8.14	−2.24	−1.17	0.64	0.51	2.65	2.95	1.65	6.97	40.70	−7.99	−1.98	−0.96	0.65	0.53	2.53	2.60
23	1.59	7.26	41.11	−8.98	−2.28	−1.35	0.40	0.59	2.92	3.23	1.81	7.19	41.32	−8.47	−1.75	−1.00	0.50	0.50	2.59	3.01
24	2.09	7.82	43.49	−9.90	−2.09	−1.59	0.51	0.35	3.13	4.02	2.04	7.57	42.26	−9.06	−1.40	−1.05	0.26	0.45	2.65	3.73
25	2.99	9.58	49.11	−11.99	−1.94	−1.84	0.62	0.20	3.89	6.56	2.64	8.93	44.86	−10.03	−0.49	−1.08	−0.31	0.26	2.72	6.11

The table summarizes the properties of a set of 25 portfolios that are formed using the final parameter estimates for the FF-MGARCH model. For each month in the sample period, I sort the stocks in ascending order of their fitted conditional excess returns, and use the resulting percentile breakpoints to form 25 groups. The first group contains the stocks that are below the 4th percentile after sorting, the second group contains the stocks that are between the 4th and 8th percentiles after sorting, and so forth. Once I have the 25 groups for a given month, I assign equal weights to the stocks contained in each of the 25 groups to obtain excess portfolio returns. The initial ten columns of the table report the sample mean and sample volatility of the excess portfolio returns, followed by their sample covariances with the market, size, value, profitability, investment, accruals, offerings, and sales factors. The final ten columns report the corresponding model-based estimates of these unconditional moments and comoments. The sample period is July 1963 to December 2016.

Table 9
Properties of 25 portfolios formed on estimated conditional covariance with the market factor

	Sample moments of portfolio returns										Estimates implied by fitted conditional moments									
			Covariances with fundamental factors										Covariances with fundamental factors							
№	Mean	Vol	f_{MAR}	f_{SIZ}	f_{VAL}	f_{PRO}	f_{INV}	f_{ACC}	f_{OFF}	f_{SAL}	Mean	Vol	f_{MAR}	f_{SIZ}	f_{VAL}	f_{PRO}	f_{INV}	f_{ACC}	f_{OFF}	f_{SAL}
1	0.58	4.15	20.13	0.33	−0.59	−0.16	0.61	0.08	0.94	1.04	0.54	5.72	24.22	3.49	−0.57	0.56	0.96	0.09	0.59	0.65
2	0.67	4.50	23.66	−0.15	−0.67	−0.14	0.65	0.13	1.11	1.27	0.63	5.53	27.62	1.49	−1.14	0.26	1.11	0.19	1.00	0.83
3	0.65	4.70	25.65	−0.68	−1.05	−0.17	0.69	0.18	1.20	1.32	0.71	5.49	29.46	0.26	−1.42	0.08	1.16	0.26	1.22	0.95
4	0.71	4.99	28.01	−1.22	−1.27	−0.22	0.77	0.26	1.38	1.52	0.77	5.51	30.84	−0.72	−1.63	−0.05	1.17	0.30	1.40	1.05
5	0.78	5.15	29.39	−1.68	−1.44	−0.18	0.82	0.26	1.44	1.49	0.82	5.57	32.00	−1.47	−1.79	−0.16	1.20	0.34	1.54	1.13
6	0.78	5.33	30.93	−2.31	−1.74	−0.34	0.84	0.36	1.59	1.61	0.87	5.63	32.99	−2.17	−1.93	−0.25	1.21	0.38	1.66	1.22
7	0.76	5.45	31.91	−2.67	−1.91	−0.31	0.95	0.40	1.70	1.57	0.91	5.71	33.90	−2.76	−2.08	−0.33	1.23	0.41	1.77	1.28
8	0.83	5.68	33.44	−3.18	−2.15	−0.43	0.99	0.45	1.84	1.44	0.94	5.80	34.74	−3.29	−2.19	−0.41	1.26	0.45	1.87	1.35
9	0.84	5.83	34.58	−3.64	−2.25	−0.47	0.97	0.51	1.96	1.64	0.97	5.89	35.53	−3.80	−2.33	−0.48	1.27	0.47	1.96	1.40
10	0.87	5.91	35.00	−3.99	−2.24	−0.57	0.97	0.48	2.01	1.75	1.00	5.98	36.27	−4.28	−2.45	−0.54	1.28	0.50	2.05	1.45
11	0.90	6.12	36.60	−4.75	−2.64	−0.65	0.98	0.57	2.23	1.68	1.03	6.09	37.00	−4.72	−2.56	−0.61	1.30	0.52	2.14	1.50
12	0.85	6.21	37.29	−4.99	−2.66	−0.70	1.01	0.59	2.26	1.73	1.05	6.20	37.74	−5.15	−2.69	−0.67	1.31	0.54	2.23	1.55
13	0.88	6.46	38.92	−5.57	−2.83	−0.77	1.29	0.59	2.40	2.01	1.07	6.31	38.50	−5.58	−2.81	−0.73	1.34	0.56	2.32	1.61
14	0.84	6.49	39.12	−5.87	−3.11	−0.87	1.25	0.58	2.42	1.83	1.09	6.43	39.24	−5.99	−2.93	−0.80	1.37	0.58	2.41	1.67
15	0.83	6.59	39.53	−6.42	−3.24	−1.02	1.28	0.62	2.55	2.17	1.10	6.56	40.01	−6.42	−3.11	−0.87	1.39	0.61	2.50	1.71
16	0.95	6.82	41.01	−6.85	−3.44	−0.97	1.35	0.61	2.66	1.99	1.11	6.70	40.83	−6.82	−3.25	−0.94	1.44	0.63	2.60	1.77
17	0.91	7.16	42.85	−7.91	−3.74	−1.38	1.36	0.66	2.94	2.24	1.12	6.86	41.69	−7.26	−3.45	−1.01	1.48	0.65	2.70	1.82
18	0.95	7.17	42.89	−8.14	−3.88	−1.19	1.38	0.69	3.04	2.20	1.11	7.02	42.62	−7.66	−3.64	−1.08	1.58	0.67	2.81	1.86
19	0.84	7.37	43.61	−8.66	−4.33	−1.51	1.42	0.77	3.09	2.07	1.10	7.21	43.66	−8.09	−3.87	−1.17	1.68	0.70	2.94	1.92
20	0.86	7.57	44.58	−9.33	−4.33	−1.54	1.48	0.77	3.32	2.31	1.08	7.44	44.85	−8.52	−4.11	−1.25	1.81	0.72	3.08	1.98
21	0.94	8.00	46.68	−10.17	−4.71	−1.88	1.64	0.70	3.56	2.59	1.03	7.70	46.27	−8.98	−4.45	−1.37	2.04	0.74	3.26	2.04
22	0.89	8.19	47.51	−10.70	−5.15	−1.88	1.65	0.68	3.76	2.46	0.94	8.05	48.08	−9.43	−4.88	−1.49	2.39	0.77	3.49	2.11
23	0.83	8.67	49.96	−10.85	−5.24	−2.15	1.62	0.70	4.03	2.77	0.83	8.56	50.55	−10.01	−5.43	−1.66	2.88	0.79	3.81	2.26
24	0.93	9.25	52.42	−11.79	−5.46	−2.47	2.09	0.62	4.55	3.20	0.67	9.36	54.29	−10.84	−6.20	−1.90	3.59	0.82	4.30	2.54
25	1.25	10.43	57.27	−14.43	−6.28	−3.15	2.04	0.73	5.30	4.75	0.56	11.25	62.50	−12.96	−7.26	−2.43	4.67	0.86	5.35	3.95

The table summarizes the properties of a set of 25 portfolios that are formed using the final parameter estimates for the FF-MGARCH model. For each month in the sample period, I sort stocks in ascending order of the fitted conditional covariance of their excess returns with the market factor, and use the resulting percentile breakpoints to form 25 groups. The first group contains the stocks that are below the 4th percentile after sorting, the second group contains the stocks that are between the 4th and 8th percentiles after sorting, and so forth. Once I have the 25 groups for a given month, I assign equal weights to the stocks contained in each of the 25 groups to obtain excess portfolio returns. The initial ten columns of the table report the sample mean and sample volatility of the excess portfolio returns, followed by their sample covariances with the market, size, value, profitability, investment, accruals, offerings, and sales factors. The final ten columns report the corresponding model-based estimates of these unconditional moments and comoments. The sample period is July 1963 to December 2016.

Table 10
Pricing performance of fundamental factors in the unconditional APT framework

	Eight fundamental factors						Five Fama-French factors (2 x 3)					
	Table 8 portfolios			Table 9 portfolios			Table 8 portfolios			Table 9 portfolios		
№	Int	t-stat	$R^2(\%)$	Int	t-stat	$R^2(\%)$	Int	t-stat	$R^2(\%)$	Int	t-stat	$R^2(\%)$
1	0.30	2.23	89.2	0.23	2.50	83.0	−0.49	−3.27	84.0	−0.01	−0.19	89.1
2	0.37	3.22	90.9	0.15	2.09	91.2	−0.13	−1.08	87.2	−0.01	−0.18	90.4
3	0.30	3.17	92.8	0.04	0.64	92.9	−0.18	−1.72	89.5	−0.07	−1.10	89.
4	0.30	3.50	93.0	0.06	0.92	94.4	−0.10	−1.08	90.5	−0.04	−0.66	91.1
5	0.17	2.55	95.1	0.04	0.80	94.9	−0.09	−1.18	92.5	0.03	0.44	91.0
6	0.17	2.39	94.8	0.04	0.61	95.0	−0.09	−1.29	92.9	0.01	0.08	91.9
7	0.12	1.89	95.6	0.02	0.41	95.0	−0.08	−1.15	93.1	−0.01	−0.12	92.1
8	0.02	0.37	95.2	0.05	0.82	94.9	−0.10	−1.35	92.3	0.06	0.88	92.2
9	0.03	0.52	95.9	0.02	0.30	95.3	−0.09	−1.38	93.8	0.07	0.93	92.3
10	−0.08	−1.34	95.7	−0.00	−0.04	94.0	−0.14	−1.97	93.2	0.09	1.07	90.9
11	−0.08	−1.27	95.4	−0.00	−0.03	95.0	−0.06	−0.65	91.3	0.15	1.79	91.5
12	−0.12	−2.14	95.6	−0.08	−1.30	95.5	−0.03	−0.44	91.9	0.08	0.98	91.7
13	−0.18	−2.78	95.0	−0.02	−0.34	95.6	−0.01	−0.14	90.6	0.13	1.22	89.9
14	−0.20	−2.93	94.6	−0.09	−1.49	95.7	0.00	0.00	90.3	0.12	1.29	89.9
15	−0.12	−1.64	94.6	−0.09	−1.24	94.7	0.11	1.14	88.4	0.11	1.09	87.9
16	−0.25	−3.88	94.2	−0.04	−0.63	95.0	0.02	0.19	87.7	0.24	2.07	87.9
17	−0.30	−4.69	94.6	−0.10	−1.22	95.2	0.05	0.43	86.	0.24	1.82	85.4
18	−0.24	−3.23	94.0	−0.05	−0.79	95.6	0.17	1.52	85.1	0.30	2.19	84.4
19	−0.24	−3.38	94.0	−0.10	−1.31	94.4	0.20	1.55	83.	0.21	1.48	83.7
20	−0.32	−4.41	93.9	−0.12	−1.37	94.3	0.19	1.57	82.3	0.30	2.10	82.5
21	−0.27	−3.62	94.3	−0.06	−0.53	93.8	0.35	2.38	79.4	0.38	2.25	78.9
22	−0.21	−2.56	93.8	−0.11	−1.15	93.9	0.50	2.97	76.1	0.35	1.99	79.
23	−0.10	−1.01	93.0	−0.17	−1.47	92.0	0.77	4.44	73.0	0.28	1.56	78.2
24	0.17	1.95	93.9	0.09	0.66	91.3	1.25	6.03	67.2	0.45	2.11	75.2
25	0.77	6.89	93.0	0.31	2.46	93.0	2.12	6.65	52.8	0.70	2.52	66.4

The table reports the results of regression-based pricing tests for the portfolios examined in Tables 8 and 9. The footnotes to these tables describe the sorting schemes used to form the portfolios. I conduct the tests by fitting time-series regressions to the excess portfolio returns. The regressors are either a constant and the eight fundamental factors that appear in the FF-MGARCH model, or a constant and the five Fama and French (2015) factors. I fit each regression by OLS, and report the estimated intercept (Int), its t-statistic (t-stat), and the regression R-squared (R^2). The t-statistics are robust to conditional heteroskedasticity. The sample period is July 1963 to December 2016.

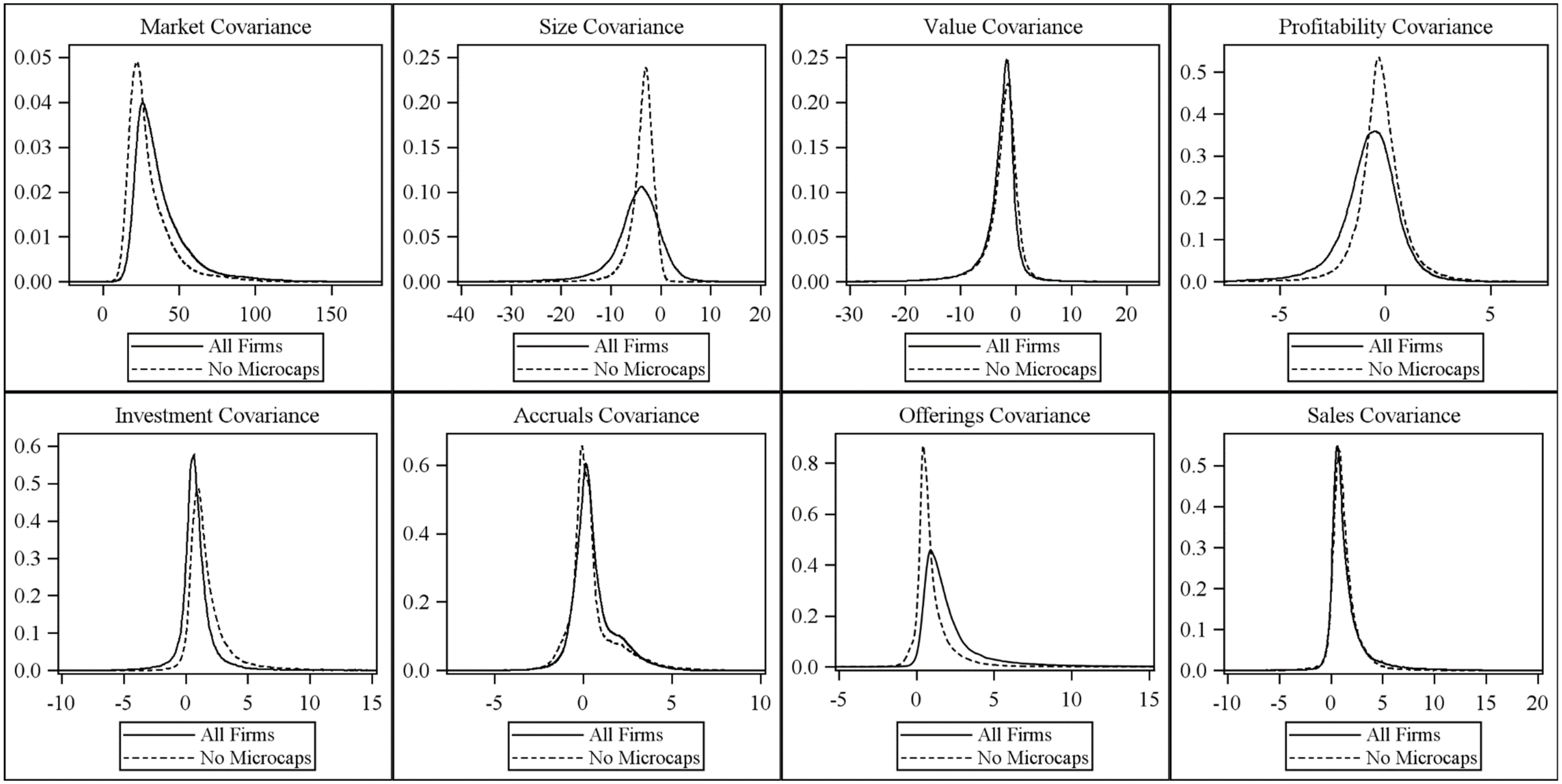

Figure 1. Estimated densities of conditional covariances between individual excess stock returns and fundamental factors. The figure shows estimated densities of the conditional covariances of individual excess stock returns with the size, value, profitability, investment, accruals, offerings, and sales factors. To create the plots, I randomly select 250 stocks each month from the set of available stocks for the month, record the fitted conditional covariances between the excess returns on the selected stocks and the factors for every month in the sample period, and use a Gaussian kernel with the bandwidth computed via the plug-in formula of Sheather and Jones (1991) to estimate the unconditional density of fitted conditional covariances with each factor. The fitted conditional covariances are based on the final parameter estimates produced by the iterative estimation procedure for the FF-MGARCH model. The solid lines show the density estimates using the full set of NYSE, AMEX, and NASDAQ firms. The dashed lines show the density estimates for a sample that excludes firms whose market equity for the month is less than the 20th percentile of the monthly cross-sectional distribution of market equity for NYSE firms.

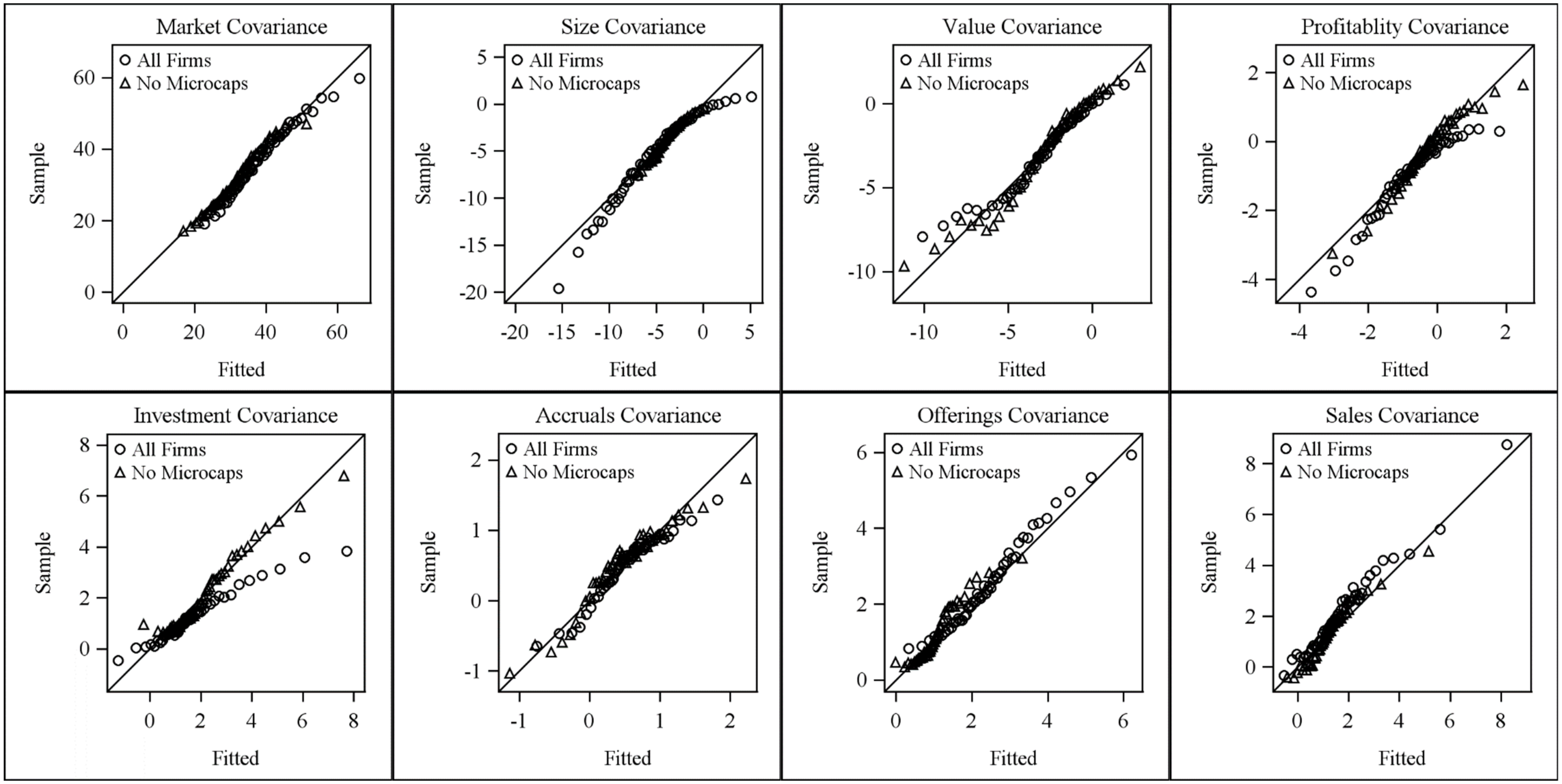

Figure 2. Sample covariances vs model-based estimates of unconditional covariances for portfolios formed on fitted conditional covariances. The figure compares the sample covariances of excess portfolio returns with the size, value, profitability, investment, accruals, offerings, and sales factors to the model-based estimates of the unconditional covariances with these factors (those implied by the fitted conditional means and fitted conditional covariances). To construct the 50 equally-weighted portfolios used for a given plot, I sort stocks on the fitted conditional covariance of their excess returns with the indicated factor for each month in the sample period. Stocks that fall below the 0.025 quantile a given month are assigned to portfolio 1, those that fall between the 0.025 and 0.05 quantiles are assigned to portfolio 2, and so on. The fitted conditional covariances are based on the final parameter estimates produced by the iterative estimation procedure for the FF-MGARCH model. The circles denote estimates for the full set of NYSE, AMEX, and NASDAQ firms. The triangles denote estimates for a sample that excludes firms whose market equity for the month is less than the 20th percentile of the monthly cross-sectional distribution of market equity for NYSE firms.

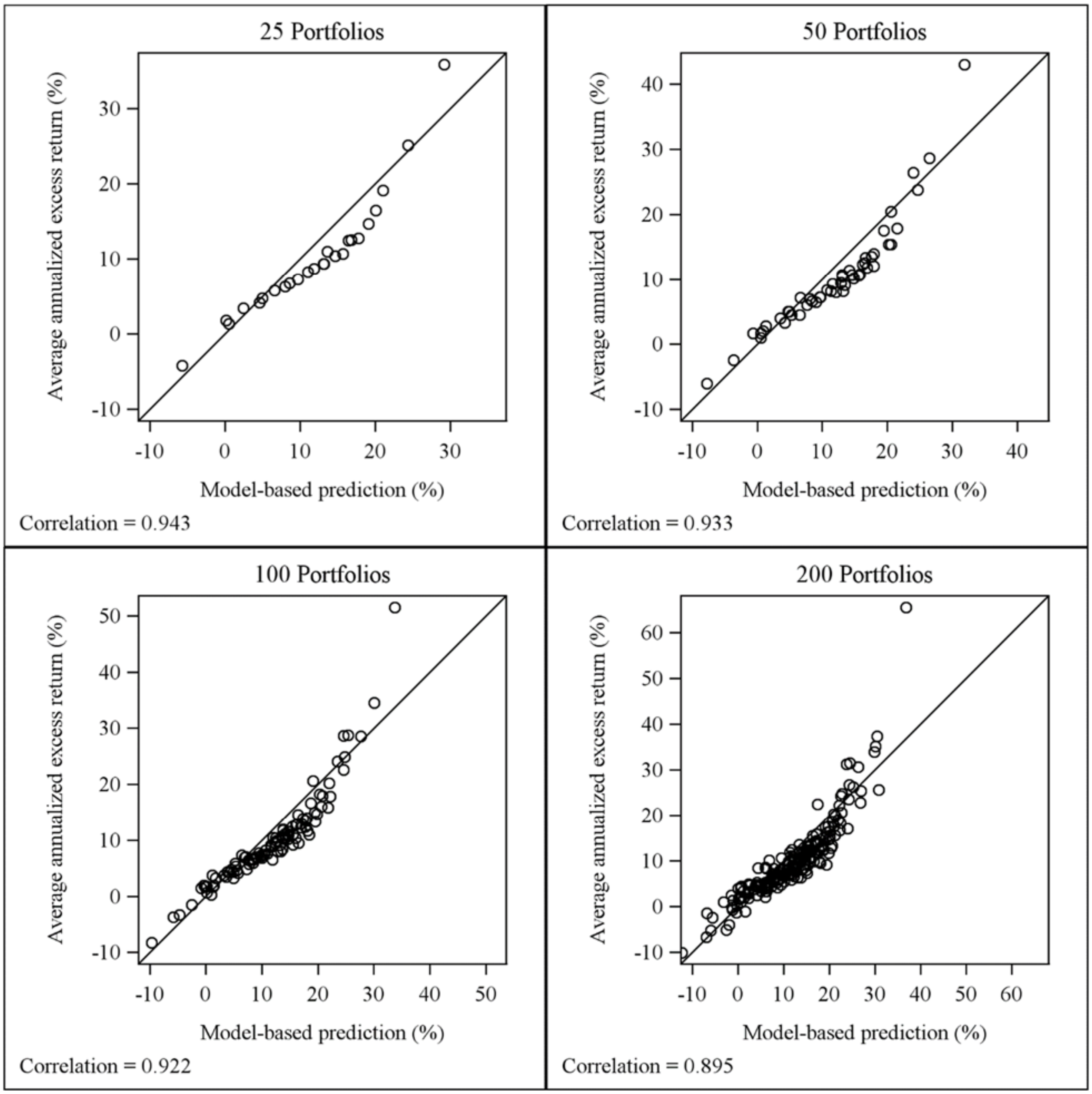

Figure 3. Average returns vs average fitted conditional expected returns for portfolios formed on fitted conditional expected returns. The figure compares average excess returns to model-based estimates of unconditional expected excess returns (those implied by the fitted conditional expected excess returns) for several sets of equally-weighted portfolios. To construct the $N \in (25, 50, 100, 200)$ portfolios used for a given plot, I sort stocks on their fitted conditional expected excess returns for each month in the sample period. Stocks that fall below the $1/N$ quantile for a given month are assigned to portfolio 1, those that fall between the $1/N$ and $2/N$ quantiles are assigned to portfolio 2, and so on. The fitted conditional expected excess returns are based on the final parameter estimates produced by the iterative estimation procedure for the FF-MGARCH model using all NYSE, AMEX, and NASDAQ firms.